

STATISTICAL HYPERBOLICITY IN TEICHMÜLLER SPACE

SPENCER DOWDALL, MOON DUCHIN, AND HOWARD MASUR

ABSTRACT. In this paper we explore the idea that Teichmüller space with the Teichmüller metric is hyperbolic “on average.” We consider several different measures on Teichmüller space and show that with respect to each one, the average distance between points in a ball of radius r is asymptotic to $2r$, which is as large as possible.

1. INTRODUCTION

Let S be a closed surface of genus $g > 1$. In this paper we continue the study of Teichmüller space $\mathcal{T}(S)$, which is the parameter space for several types of geometric structures on S . It is known that Teichmüller space equipped with the Teichmüller metric $d_{\mathcal{T}}$ is a complete metric space homeomorphic to \mathbb{R}^{6g-6} . It is not δ -hyperbolic [16], and several kinds of obstructions to hyperbolicity are known: for instance, pairs of geodesic rays through the same point may fellow-travel arbitrarily far apart [10], and there are large “thin parts” of the space which, up to bounded additive error, are isometric to product spaces equipped with sup metrics [17] (and therefore not hyperbolic). These exceptions to the negative-curvature phenomena seem to come from rare occurrences, so one might expect properties that are characteristic of hyperbolicity to hold on average.

One way to make this precise is to consider the average distance between points on metric spheres in a metric space (X, d) . Writing $\mathcal{S}_r(x)$ for the sphere of radius r based at x , we define a geometric statistic for the large spheres as follows. Given a family of probability measures μ_r on the spheres $\mathcal{S}_r(x)$, let $E(X) = E(X, x, d, \{\mu_r\})$ be the average distance between points on large spheres:

$$E(X) := \lim_{r \rightarrow \infty} \frac{1}{r} \int_{\mathcal{S}_r(x) \times \mathcal{S}_r(x)} d(y, z) \, d\mu_r(y) d\mu_r(z),$$

if the limit exists. It is shown in [6] that non-elementary hyperbolic groups all have $E(G, S) = 2$ for any finite generating set S ; this is also the case in the hyperbolic space \mathbb{H}^n of any dimension endowed with the natural measure on spheres. By contrast, it is shown that $E(\mathbb{R}^n) < 2$ (increasing over the range $[4/\pi, \sqrt{2})$ as the dimension goes from 2 to infinity), and that $E(\mathbb{Z}^n, S) < 2$ for all n and S , with nontrivial dependence on S . (See [6] for more examples.)

In the case of Teichmüller space, one could consider a number of a priori different measures that are natural from various points of view. For instance, as a metric space $\mathcal{T}(S)$ carries a $(6g-6)$ -dimensional *Hausdorff measure* η . Other measures come from the Finsler structure (Busemann measure and Holmes-Thompson

Date: May 1, 2012.

The second author was partially supported by NSF DMS-0906086. The third author was partially supported by NSF DMS-0905907.

measure), from the holonomy coordinates on the cotangent bundle (holonomy, or Masur-Veech, measure), and from the symplectic structure. In §3, we find that all of these are absolutely continuous with respect to each other, and in fact are bounded in terms of each other. Furthermore, the sphere $\mathcal{S}_r(x)$ can be identified with the unit sphere $\mathcal{Q}^1(x)$ in the vector space of quadratic differentials on x via the Teichmüller map. The latter has various natural measures, and corresponding measures on $\mathcal{S}_r(x)$ will be called *visual measures*; we will pay special attention to two standard visual measures, denoted $\text{Vis}(\nu_x)$ and $\text{Vis}(s_x)$.

Our main theorem concerns the average distance between points in the ball $\mathcal{B}_r(x)$ of radius r centered at x . We seek to show that the average distance between points in $\mathcal{B}_r(x)$ is asymptotic to $2r$, which, in light of the triangle inequality, is the maximum possible distance. Two axioms are needed for this conclusion: a measure must satisfy a *thickness estimate*, guaranteeing that typical rays spend a definite proportion of their time in the thick part, and an *exponential decay estimate*, asserting that fellow-traveling becomes exponentially rare. All of the measures discussed above satisfy these axioms.

Theorem 1. *Suppose μ satisfies a thickness estimate and an exponential decay estimate. Then for every point $x \in \mathcal{T}(S)$,*

$$\lim_{r \rightarrow \infty} \frac{1}{r} \frac{1}{\mu(\mathcal{B}_r(x))^2} \int_{\mathcal{B}_r(x) \times \mathcal{B}_r(x)} d_{\mathcal{T}}(y, z) \, d\mu(y) d\mu(z) = 2.$$

With respect to the standard visual measures, the same methods produce the following result.

Theorem 2. *$E(\mathcal{T}(S)) = 2$ with respect to Teichmüller distance and the standard visual measures.*

We sketch here the main ideas in the proofs. The first step is to show (§5.2) that most pairs of geodesics separate from each other in the Teichmüller metric after a threshold time. Then one would hope that, as in a hyperbolic space, the geodesic joining their endpoints would follow the first geodesic back to approximately where they separate before following the other so that its length is roughly the sum of the lengths of the two geodesics, as on the left in Figure 1. The Minsky product

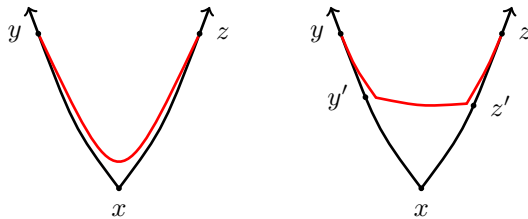


FIGURE 1. We will show that the geodesic between points on generic rays “dips” back near the basepoint. This requires an analysis of the time spent in thick and thin parts: if $[x, y']$ and $[x, z']$ lie in thin parts corresponding to disjoint subsurfaces, then Minsky’s product regions theorem shows that the connecting geodesic can take a “shortcut.” We show that this effect is rare.

regions theorem [17] says that this in fact may not happen. If the pair of geodesics

separate because they enter thin parts corresponding to disjoint subsurfaces, then the geodesic joining their endpoints travels through those thin regions simultaneously and its length is then smaller than the sum, as on the right in Figure 1. Our goal is to show that this phenomenon does not happen on average. The mechanism for showing this is the coarsely contracting map from $\mathcal{T}(S)$ to the curve complex. Using thickness statistics for the rays, we show (Theorem 40) that most pairs of Teichmüller geodesics separate in the curve complex after a bounded time, which rules out shortcuts through product regions. We then show (Theorem 41) that for most pairs, a geodesic connecting the two rays must pass through the region where the rays separated, which immediately provides the needed bounds.

Along the way we use a wide array of tools in Teichmüller geometry, including some that are quite recently developed: subsurface projection and reverse triangle inequalities (following Masur–Minsky and Rafi), time-ordering of thin intervals along Teichmüller geodesics (Rafi), antichain bounds (Rafi and Schleimer); volume asymptotics in Teichmüller space (Athreya–Bufetov–Eskin–Mirzakhani) and random walks as discretized Teichmüller geodesics (Eskin–Mirzakhani). We also develop several facts that can serve as future tools, including a simplified formula for Teichmüller distance (§2.4), comparisons of measures (§3), and the results showing that thickness estimates for rays ensure that a typical ray makes definite progress in the curve complex and that typical pairs obtain big distance in the curve complex and big distance in Teichmüller space (§§6.1–6.2–6.3).

1.1. Acknowledgments. We would like to thank Alex Eskin, Benson Farb, Curtis McMullen, and especially Kasra Rafi for numerous helpful comments and explanations.

2. BACKGROUND MATERIAL

2.1. Teichmüller space and quadratic differentials. Recall that Teichmüller space $\mathcal{T}(S)$ is the space of marked Riemann surfaces X that are homeomorphic to the topological surface S . More precisely, it consists of pairs (X, f) , where $f : S \rightarrow X$ is a homeomorphism, up to the equivalence relation that $(X_1, f_1) \sim (X_2, f_2)$ when there exists a conformal map $F : X_1 \rightarrow X_2$ such that $F \circ f_1$ is isotopic to f_2 . Alternately, we may define $\mathcal{T}(S)$ as the space of marked hyperbolic surfaces (ρ, f) ; namely, maps $f : S \rightarrow \rho$ with $(\rho_1, f_1) \sim (\rho_2, f_2)$ when there exists an isometry $F : \rho_1 \rightarrow \rho_2$ s.t. $F \circ f_1$ is isotopic to f_2 .

Using the first definition of $\mathcal{T}(S)$, the Teichmüller distance is given by

$$d_{\mathcal{T}}((X_1, f_1), (X_2, f_2)) := \inf_{F \sim f_2 \circ f_1^{-1}} \frac{1}{2} \log K(F),$$

where the minimum is taken over all quasiconformal maps F and $K(F)$ is the maximal dilatation of F . The space $\mathcal{T}(S)$ is homeomorphic to the ball \mathbb{R}^{6g-6} , and from now on we will use $h = 6g - 6$ to designate this dimension. In this paper, we will denote a point of $\mathcal{T}(S)$ by x , regarding it either as a Riemann surface or a hyperbolic surface, and suppressing the marking f .

For $x, y \in \mathcal{T}(S)$, the Teichmüller geodesic segment joining x to y will usually be denoted $[x, y]$. We will also use $\gamma(t)$ to denote a geodesic ray or segment when the time parameter is important.

A *quadratic differential* on a Riemann surface X is a holomorphic 2-tensor $q = \phi(z)dz^2$ on X . The space of all quadratic differentials on all Riemann surfaces

homeomorphic to S is denoted $\mathcal{Q}(S)$. A point of $\mathcal{Q}(S)$ will be denoted q , with the underlying complex structure implicit in the notation. The real dimension of $\mathcal{Q}(S)$ is $12g - 12 = 2h$. Reading off the Riemann surface, we obtain a projection to the Teichmüller space $\pi : \mathcal{Q}(S) \rightarrow \mathcal{T}(S)$. Under this projection, $\mathcal{Q}(S)$ forms vector bundle over $\mathcal{T}(S)$ which is canonically identified with the cotangent bundle of $\mathcal{T}(S)$. Each fiber $\mathcal{Q}(X)$ is equipped with a norm given by the total area of q ; namely $\|q\| = \int_X |\phi(z)dz^2|$. Recall that $d_{\mathcal{T}}$ is not a Riemannian metric on $\mathcal{T}(S)$, but rather a Finsler metric; it comes from dualizing the norm on \mathcal{Q} to give a norm on each tangent space of $\mathcal{T}(S)$ that is not induced by any inner product.

It is the famous theorem of Teichmüller that the infimum in the definition of $d_{\mathcal{T}}$ is realized uniquely by a *Teichmüller map* from X_1 to X_2 . A Teichmüller map is determined by an initial quadratic differential $q = \phi(z)dz^2$ on X_1 and the number K . The Teichmüller map expands along the horizontal trajectories of q by a factor of $K^{1/2}$ and contracts along the vertical trajectories by the same factor to obtain a terminal quadratic differential q' on the image surface X_2 . If we fix q and let $K = e^{2t}$ vary over $t \in [0, \infty)$ we get a Teichmüller geodesic ray.

Recall that the *mapping class group* of S , defined by

$$\text{Mod}(S) := \text{Diff}^+(S)/\text{Diff}_0(S),$$

is the discrete group of orientation-preserving diffeomorphisms of S , up to isotopy. This group acts isometrically on $\mathcal{T}(S)$ by changing the marking: $\phi \cdot (X, f) = (X, f \circ \phi^{-1})$. In fact, by a result of Royden [21], $\text{Mod}(S)$ is the full group of (orientation-preserving) isometries of $(\mathcal{T}(S), d_{\mathcal{T}})$.

2.2. Curve graph. When we speak of a *curve* on S , this will mean an isotopy class of essential simple closed curves. Given $x \in \mathcal{T}(S)$, the *length* $l_x(\alpha)$ of a curve α is the length of the geodesic in the isotopy class in the hyperbolic metric x .

We recall the definition of the *curve complex* (or curve graph) $\mathcal{C}(S)$ of S . The vertices of $\mathcal{C}(S)$ are curves on S . Two vertices are joined by an edge if the corresponding curves can be realized disjointly. Assigning edges to have length 1 we have a metric graph. Properly speaking, $\mathcal{C}(S)$ is the flag complex associated to this curve graph, but since we are working coarsely, we can identify $\mathcal{C}(S)$ with the graph.

It is known that the curve graph is hyperbolic [15]. That is, there exists a constant $\delta > 0$ such that every geodesic triangle in $\mathcal{C}(S)$ is δ -thin: each side of the triangle is contained in the union of the δ -neighborhoods of the other two sides. It follows that every geodesic quadrilateral in $\mathcal{C}(S)$ is 2δ -thin (each side is within 2δ of the union of the other three sides). Furthermore, in any δ -hyperbolic metric space and for any quasi-isometry constants (K, C) , there exists a constant τ , depending only on δ, K, C , such that any two (K, C) -quasi-geodesic segments with the same endpoints remain within τ of each other. Since actual geodesics are $(1, 0)$ -quasi-geodesics, this implies that every (K, C) -quasi-geodesic quadrilateral is $2(\delta + \tau)$ -thin.

2.3. Thick parts and subsurface projections. For any given ϵ_0 , we say a curve is ϵ_0 -short if its hyperbolic length is less than ϵ_0 . Then define \mathcal{T}_{ϵ_0} , the ϵ_0 -*thick part* of Teichmüller space, to be the subset of $x \in \mathcal{T}(S)$ on which no curve is ϵ_0 -short.

For each $x \in \mathcal{T}(S)$ there is associated a *Bers marking* μ_x . To construct μ_x , greedily choose a shortest *pants decomposition* of the surface (a collection of $3g - 3$ disjoint simple geodesics). Then for each pants curve β , choose a shortest geodesic

crossing β minimally (either once or twice depending on the topology) that is disjoint from all other pants curves. The total collection of $6g - 6$ curves is called a Bers marking and is defined up to finitely many choices.

Throughout, a *proper subsurface of S* will mean a compact, properly embedded subsurface $V \subset S$ which is not equal to S and for which the induced map on fundamental groups is injective. Subsurfaces which are isotopic to each other will not be considered distinct. The proper subsurfaces of S fall into two categories, annuli and non-annuli, which behave somewhat differently. Nevertheless, we will strive to develop intuitive notation under which these two possibilities may be dealt with on equal footing.

Every proper subsurface V has a nonempty boundary ∂V consisting of a disjoint union of curves on S . We say that two subsurfaces V and W *transversely intersect*, denoted $V \pitchfork W$, if they are neither (isotopically) disjoint nor nested. In this case, ∂V necessarily intersects W , and ∂W intersects V .

Consider a non-annular subsurface V , possibly equal to S . The *subsurface projection* $\pi_V(\beta)$ of a simple closed curve $\beta \subset S$ to V is defined as follows: Realize β and ∂V as geodesics (in any hyperbolic metric on S). If $\beta \subset V$, then $\pi_V(\beta)$ is defined to be β . If β is disjoint from V , then $\pi_V(\beta)$ is undefined. Otherwise, $\beta \cap V$ is a disjoint union of finitely many homotopy classes of arcs with endpoints on ∂V , and we obtain $\pi_V(\beta)$ by choosing any arc and performing a surgery along ∂V to create a simple closed curve contained in V . The subsurface projection of a point $x \in \mathcal{T}(S)$ is then defined to be the collection

$$\pi_V(x) := \{\pi_V(\beta)\}_{\beta \in \mu_x}$$

of curves obtained by varying β in the Bers marking at x . This is a non-empty subset of the curve complex $\mathcal{C}(V)$ with uniformly bounded diameter.

Definition 3 (Non-annular projection distance). For a non-annular subsurface $V \subseteq S$, the *projection distance in V* of a pair of points $x, y \in \mathcal{T}(S)$ is defined to be

$$d_V(x, y) := \text{diam}_{\mathcal{C}(V)}(\pi_V(x) \cup \pi_V(y)).$$

In particular, $d_S(x, y)$ denotes the curve complex distance. When convenient, we will also denote this distance by $d_{\mathcal{C}(V)} := d_V$.

For an annular subsurface $A \subset S$ with *core curve* $\alpha = \partial A$, there are two kinds of projection distances: one that measures twisting about α and is analogous to the definition above, and a second which also incorporates the length of α . Any simple closed curve β that crosses α may be realized by a geodesic and then lifted to a geodesic $\tilde{\beta}$ in the annular cover \tilde{A} , that is, the quotient of \mathbb{H}^2 by the deck transformation corresponding to α , with the Gromov compactification. For a pair β, γ of such curves, we may then consider the intersection number $i(\tilde{\beta}, \tilde{\gamma})$ in \tilde{A} . The *twisting distance in A* of a pair of points $x, y \in \mathcal{T}(S)$ is then defined as

$$d_{\mathcal{C}(A)}(x, y) := \sup_{\beta \in \mu_x, \gamma \in \mu_y} i_{\tilde{A}}(\tilde{\beta}, \tilde{\gamma}).$$

We additionally define a hyperbolic projection distance as follows.

Definition 4 (Annular projection distance). For an annular subsurface $A \subset S$ with core curve $\alpha = \partial A$, denote by \mathbb{H}_α a copy of the standard horoball $\{\text{Im}(z) \geq 1\} \subset \mathbb{H}^2$. Given $x, y \in \mathcal{T}(S)$, we consider the points $(0, 1/l_x(\alpha))$ and $(d_{\mathcal{C}(A)}(x, y), 1/l_y(\alpha)) \in$

\mathbb{H}^2 and denote their closest point projections to the horoball \mathbb{H}_α by

$$\pi_\alpha(x) = \left(0, \max \left\{1, \frac{1}{l_x(\alpha)}\right\}\right), \quad \pi_\alpha(y) = \left(d_{\mathcal{C}(A)}(x, y), \max \left\{1, \frac{1}{l_y(\alpha)}\right\}\right).$$

The *projection distance* in A (or *hyperbolic distance* $d_{\mathbb{H}_\alpha}$) between x and y is then defined to be

$$d_A(x, y) := d_{\mathbb{H}^2}(\pi_\alpha(x), \pi_\alpha(y)).$$

2.4. Distance formula. For functions f, g and constants $K \geq 1, C \geq 0$, we will use the notation $f(x) \stackrel{K, C}{\asymp} g(x)$ if the inequalities $\frac{1}{K}g(x) - C \leq f(x) \leq K \cdot g(x) + C$ hold for all x . As usual we denote $f = O(g)$ if the second inequality holds, and $f = o(g)$ if $f(x)/g(x) \rightarrow 0$ as $x \rightarrow \infty$.

The following distance formula due to Rafi relates the Teichmüller distance between two points x and y to the combinatorics of the corresponding Bers markings μ_x and μ_y .

Theorem 5 (Distance formula, Rafi [18]). *Fix a small $\epsilon_0 > 0$. For any sufficiently large threshold M_0 , there exist quasi-isometry constants $K \geq 1$ and $C \geq 0$ depending only on M_0 and the topology of S such that, for all $x, y \in \mathcal{T}(S)$ we have*

$$\begin{aligned} d_{\mathcal{T}}(x, y) &\stackrel{K, C}{\asymp} d_S(x, y) + \sum_V [d_V(x, y)]_{M_0} + \max_{\alpha \in \Gamma_{xy}} d_{\mathbb{H}_\alpha}(x, y) \\ &\quad + \sum_{A: \partial A \not\subset \Gamma_{xy}} \log_+[d_{\mathcal{C}(A)}(x, y)]_{M_0} + \max_{\alpha \in \Gamma_x} \log_+\left(\frac{1}{l_x(\alpha)}\right) + \max_{\alpha \in \Gamma_y} \log_+\left(\frac{1}{l_y(\alpha)}\right), \end{aligned}$$

where Γ_{xy} is the set of ϵ_0 -short curves in both x and y , Γ_x is the set of curves that are ϵ_0 -short in x but not in y , and Γ_y is defined similarly. Here and throughout, \log_+ is a modified logarithm so that $\log_+ a = 0$ for $a \in [0, 1]$, and $[\cdot]_{M_0}$ is a threshold function for which $[N]_{M_0} := N$ when $N \geq M_0$ and $[N]_{M_0} := 0$ otherwise.

We will make all annular measurements with the hyperbolic distance on the \mathbb{H}_α , obtaining a particularly simple restatement of this formula.

Proposition 6 (Repackaged distance formula). *With the same hypotheses as above, we have:*

$$(1) \quad d_{\mathcal{T}}(x, y) \stackrel{K, C}{\asymp} d_S(x, y) + \sum_Y [d_Y(x, y)]_{M_0}$$

Here, the sum is over all (annular and non-annular) proper subsurfaces.

Remark 7. The definition of $d_{\mathbb{H}_\alpha} = d_A$ given above is technically different than that used by Rafi in [18]; however, the two definitions agree up to bounded additive error.

This version of the distance formula treats annular and non-annular subsurfaces on equal footing. For simplicity and without loss of generality, below we suppose that ϵ_0 is fixed small enough that $\log_+(1/\epsilon_0) \geq 100$, say. We begin with a straightforward reformulation.

Lemma 8. *Given any sufficiently large threshold M_0 , there exist $K \geq 1$ and $C \geq 0$ such that for all $x, y \in \mathcal{T}(S)$ we have:*

$$d_{\mathcal{T}}(x, y) \stackrel{K, C}{\asymp} d_S(x, y) + \sum_V [d_V(x, y)]_{M_0} + \sum_{A: \partial A \in \Gamma_{xy}} [d_A(x, y)]_{M_0} + \sum_{A: \partial A \notin \Gamma_{xy}} \left[\max \left\{ \log_+(d_{\mathcal{C}(A)}(x, y)), \log_+\left(\frac{1}{l_x(\partial A)}\right), \log_+\left(\frac{1}{l_y(\partial A)}\right) \right\} \right]_{\log M_0}$$

Proof. Since Γ_{xy} , Γ_x and Γ_y each contain at most $3g - 3$ curves, each max over these sets is within bounded multiplicative error of the corresponding sum, and applying a threshold only creates bounded additive error, so the first three terms of the lemma are established. By the definition of Γ_x we have

$$\sum_{\alpha \in \Gamma_x} \log_+\left(\frac{1}{l_x(\alpha)}\right) = \sum_{\alpha \notin \Gamma_{xy}} \log_+\left[\frac{1}{l_x(\alpha)}\right]_{1/\epsilon_0}.$$

Since this is a sum with at most $3g - 3$ nonzero terms, we can increase the threshold to any number $M_0 \geq 1/\epsilon_0$ with bounded additive error. Finally, for functions f, g, h , we have

$$\log_+[f]_{M_0} + \log_+[g]_{M_0} + \log_+[h]_{M_0} \stackrel{3,0}{\asymp} [\max\{\log_+f, \log_+g, \log_+h\}]_{\log M_0}. \quad \square$$

We now show that each term in the last summand is bilipschitz equivalent to the corresponding hyperbolic distance $d_A(x, y)$.

Lemma 9. *Consider an annular subsurface $A \subset S$ with core curve $\partial A = \alpha$. For each pair of points $x, y \in \mathcal{T}(S)$, set*

$$H_A(x, y) := \max \left\{ \log_+(d_{\mathcal{C}(A)}(x, y)), \log_+\left(\frac{1}{l_x(\alpha)}\right), \log_+\left(\frac{1}{l_y(\alpha)}\right) \right\}.$$

If $x, y \in \mathcal{T}(S)$ are such that $\alpha \notin \Gamma_{xy}$ and either $d_A(x, y)$ or $H_A(x, y)$ is greater than $36 \log_+(1/\epsilon_0)$, then $6^{-1}d_A(x, y) \leq H_A(x, y) \leq 6d_A(x, y)$.

Proof. Choose points $x, y \in \mathcal{T}(S)$ that satisfy the hypotheses. To fix notation, set $\pi'_\alpha(x) = (0, 1)$ and $\pi'_\alpha(y) = (d_{\mathcal{C}(A)}(x, y), 1)$. These are the closest-point projections of $\pi_\alpha(x)$ and $\pi_\alpha(y)$ to the horocycle bounding \mathbb{H}_α , and their distances from these points are exactly given by $\log_+(1/l_x(\alpha))$ and $\log_+(1/l_y(\alpha))$. Let

$$B = d_{\mathbb{H}^2}(\pi'_\alpha(x), \pi'_\alpha(y)) = \operatorname{arccosh} \left(1 + \frac{d_{\mathcal{C}(A)}(x, y)^2}{2} \right)$$

denote the hyperbolic distance between these projections. Using this formula, one may easily check that the inequalities

$$(2) \quad \log_+d_{\mathcal{C}(A)}(x, y) \leq B \leq 4 \log_+d_{\mathcal{C}(A)}(x, y)$$

hold provided that either $B \geq 3$ or $d_{\mathcal{C}(A)}(x, y) \geq 3$.

Applying the triangle inequality with the points $\pi'_\alpha(x)$ and $\pi'_\alpha(y)$ implies that

$$(3) \quad d_A(x, y) \leq \log_+\left(\frac{1}{l_x(\alpha)}\right) + B + \log_+\left(\frac{1}{l_y(\alpha)}\right).$$

Then (2), (3), and the definition of H_A imply that $d_A(x, y) \leq 6H_A(x, y)$ in the case that $B \geq 3$. If $B < 3$, we claim that the hypotheses of the Lemma ensure that B cannot be the largest term on the right-hand side and therefore that $d_A(x, y) \leq$

$3L \leq 3H_A(x, y)$, where L denotes the larger of the other two terms. Indeed, if B were the largest term and $B < 3$, then (3) would imply $d_A(x, y) < 9$, and (2) would necessitate $\log_+ d_{C(A)}(x, y) < 3$ so that $H_A(x, y) < 3$. But then both d_A and H_A are less than 9, contradicting the hypothesis.

By the above, the assumption $d_A(x, y) \geq 36 \log_+(1/\epsilon_0)$ implies that $H_A(x, y) \geq 6 \log_+(1/\epsilon_0)$; therefore all cases will be covered by proving that this in turn implies $H_A(x, y) \leq 6d_A(x, y)$. Without loss of generality, we may assume that $l_x(\alpha) \leq l_y(\alpha)$; since $\alpha \notin \Gamma_{xy}$ this guarantees $l_y(\alpha) \geq \epsilon_0$. First suppose that $\log_+ d_{C(A)}(x, y) \geq 3 \log_+(1/l_x(\alpha))$, in which case we have $\log_+ d_{C(A)}(x, y) = H_A(x, y) \geq 6 \log_+(1/\epsilon_0)$. In particular we certainly have $d_{C(A)}(x, y) \geq 3$; thus (2) and the triangle inequality give

$$\log_+ d_{C(A)}(x, y) \leq B \leq \log_+ \left(\frac{1}{l_x(\alpha)} \right) + d_A(x, y) + \log_+ \left(\frac{1}{l_y(\alpha)} \right).$$

Therefore $H_A(x, y) = \log_+ d_{C(A)}(x, y) \leq 3d_A(x, y)$ in this case. The remaining possibility $\log_+ d_{C(A)}(x, y) \leq 3 \log_+(1/l_x(\alpha))$ necessitates $3 \log_+(1/l_x(\alpha)) \geq H_A(x, y)$. Recall that $\pi'_\alpha(x)$ is the *closest* point projection of $\pi_\alpha(x)$ to the horocycle bounding \mathbb{H}_α ; since $\pi'_\alpha(y)$ is also on this horocycle we have

$$\log_+ \left(\frac{1}{l_x(\alpha)} \right) \leq d_{\mathbb{H}^2}(\pi_\alpha(x), \pi'_\alpha(y)) \leq d_A(x, y) + \log_+ \left(\frac{1}{l_y(\alpha)} \right).$$

The assumptions $3 \log_+(1/l_x(\alpha)) \geq H_A(x, y) \geq 6 \log_+(1/\epsilon_0)$ and $l_y(\alpha) \geq \epsilon_0$ now ensure that $H_A(x, y) \leq 6d_A(x, y)$. \square

Corollary 10. *Let $H_A(x, y)$ be defined as in Lemma 9. Then for any threshold $M_0 \geq 36 \log_+(1/\epsilon_0)$ and any $x, y \in \mathcal{T}(S)$ we have*

$$\sum_{\partial A \notin \Gamma_{xy}} 6^{-1} [d_A(x, y)]_{6M_0} \leq \sum_{\partial A \notin \Gamma_{xy}} [H_A(x, y)]_{M_0} \leq \sum_{\partial A \notin \Gamma_{xy}} 6 [d_A(x, y)]_{M_0/6}$$

With these estimates, we can derive the simplified distance formula.

Proof of Repackaged Distance Formula. Fix a small $\epsilon_0 > 0$ and choose any sufficiently large threshold M_0 such that Lemma 8 holds for both e^{6M_0} and $M_0/6$ and such that $M_0/6 \geq 36 \log_+(1/\epsilon_0)$. Let $K \geq 1$ and $C \geq 0$ denote the larger of the quasi-isometry constants given by Lemma 8 for the thresholds e^{6M_0} and $M_0/6$.

Notice that, in any sum of the form $\sum [f]_M$, raising the threshold can only decrease the value of the sum, and lowering the threshold can only increase its value. Therefore, combining Lemma 8 and Corollary 10 we find that for any $x, y \in \mathcal{T}(S)$ the various distances satisfy

$$\begin{aligned} d_{\mathcal{T}} &\leq K \left(d_S + \sum_V [d_V]_{e^{6M_0}} + \sum_{\partial A \in \Gamma_{xy}} [d_A]_{e^{6M_0}} + \sum_{\partial A \notin \Gamma_{xy}} [H_A]_{6M_0} \right) + C \\ &\leq 6K \left(d_S + \sum_V [d_V]_{M_0} + \sum_{\partial A \in \Gamma_{xy}} [d_A]_{M_0} + \sum_{\partial A \notin \Gamma_{xy}} [d_A]_{M_0} \right) + C, \end{aligned}$$

where we have suppressed the x and y in the notation. The lower bound on $d_{\mathcal{T}}(x, y)$ is similar. \square

2.5. Thinness and time-ordering. We will use some results from Rafi's work [18, Prop 3.7] combinatorializing the Teichmüller metric. For every Teichmüller geodesic and every proper subsurface V , there is a (possibly empty) interval along the geodesic where ∂V is short. Outside of this interval, the projections d_V move by at most a bounded amount. In the form that we will use below: there is a global constant M and constants $\epsilon_0 < \epsilon_1$ such that for any pair of points $x, y \in \mathcal{T}(S)$ there is a possibly nonempty connected interval T_V along the geodesic segment $[x, y]$ such that

- for $a \in T_V$ the length of ∂V on a is at most ϵ_1 ;
- for $a \in [x, y] \setminus T_V$ the length of ∂V on a is at least ϵ_0 ;
- for a, b in the same component of $[x, y] \setminus T_V$, we have $d_V(a, b) < M$; and
- if $V \pitchfork W$ then $T_V \cap T_W = \emptyset$.

We should note that the interval T_V is not uniquely defined. We also note that the second condition says that on the complement of the union of thin intervals the point lies in the ϵ_0 -thick part of $\mathcal{T}(S)$. If $T_V \neq \emptyset$ we will say that V is *thin* along T_V . In particular if $d_V(x, y) \geq M$, then the interval $T_V \neq \emptyset$.

We will write $T_V < T_W$ along $[x, y]$ if both endpoints of T_V occur before both endpoints of T_W when traveling from x to y . This T_V is called the *thin interval* for V (or the *active interval* in some papers). Note that for us thin intervals are segments in Teichmüller space, whereas Rafi works with the corresponding time intervals $I_V \subset \mathbb{R}$. The geodesic $[x, y]$ is suppressed in the notation T_V , and so the same notation re-occurs when there are multiple segments in an argument; the geodesic with respect to which the interval is defined should be clear from context. We will take M to be large enough to be a valid threshold in the distance formula (1).

The properties of thin intervals imply the following *time-ordering principle*.

Lemma 11 (Time ordering [19]). *Choose any constant $M_0 \geq M$. Consider a pair of geodesic segments in \mathcal{T} with a common basepoint x which end in y, y' , respectively. Suppose $V \pitchfork W$ are transversely intersecting subsurfaces of S .*

- (1) *If $d_V(x, y), d_V(x, y'), d_W(x, y), d_W(x, y') \geq 3M_0$, then the thin intervals T_V, T_W appear in the same order along both geodesics, as they are traced out from x .*
- (2) *If $d_V(x, y), d_W(x, y), d_W(x, y') \geq 3M_0$ and T_V appears before T_W along $[x, y]$, then V determines a thin interval along $[x, y']$ which appears before T_W .*

Proof. Assume the pairs of endpoints have large projection to V and W as in the hypothesis of the first statement, and suppose T_V appears before T_W along $[x, y]$. Since the endpoints of T_V contain ∂V in their markings, we have $d_W(x, \partial V) \leq M_0$. If the intervals appear in the opposite order along $[x, y']$, then letting z be the endpoint of T_V closest to x , since z contains ∂V in its marking, we use the triangle inequality to get

$$d_W(x, \partial V) = d_W(x, z) \geq 3M_0 - M_0 = 2M_0,$$

a contradiction. This proves the first statement. Turning to the second statement, the assumption on $[x, y]$ gives us that

$$d_V(x, \partial W) \geq 2M_0.$$

Let z' be the endpoint of T_W along $[x, y']$ closest to x . It contains ∂W in its marking. We therefore have

$$d_V(x, z') \geq 2M_0 \geq M,$$

and so T_V must appear between x and z' along $[x, y']$. \square

2.6. Reverse triangle inequality. We will repeatedly use the fact that the projection of a Teichmüller geodesic to the curve complex of any subsurface forms an unparameterized quasi-geodesic that, in particular, does not backtrack. This phenomenon is captured by the following “reverse triangle inequality,” which was proved first in the case of the curve complex of the whole surface by Masur–Minsky [15] and then for general subsurfaces by Rafi [19, Thm B].

Lemma 12 (Reverse triangle inequality). *There exists $B > 0$ such that for any nonannular subsurface V (including S itself) and for any geodesic interval $[x, y]$ and any subinterval $[a, b] \subset [x, y]$ we have*

$$(4) \quad \begin{aligned} d_V(x, a) + d_V(a, y) &\leq d_V(x, y) + B, \text{ and} \\ d_V(a, b) &\leq d_V(x, y) + B. \end{aligned}$$

For an annulus A , these inequalities hold with the twisting distance $d_{C(A)}$, but not necessarily with the projection distance d_A .

In the exceptional annulus case, we have the following theorem from Rafi [18].

Theorem 13 (R.T.I. exception). *For any sufficiently large M_0 , there exists a constant $B' > 0$ with the following property. For any geodesic segment $[x, y]$ and any annulus A , if $a \in [x, y]$ is such that $d_A(x, a) + d_A(a, y) - d_A(x, y) \geq B'$ (i.e., the reverse triangle inequality fails), then there exists a proper subsurface $V \neq A$ containing a family of subsurfaces $W_i \subset V$ such that*

- $\partial V = \partial A$,
- the W_i fill V ,
- for each W_i the reverse triangle inequality (4) holds along $[x, y]$,
- $d_{W_i}(a, y) \geq M_0$ for each i ,
- $d_A(a, y) \leq \sum d_{W_i}(a, y)$.

We remark that this is not exactly how the result in [18] is stated. Rafi finds an annulus about the short curve which with respect to the quadratic differential is a disjoint union of a *flat* annulus and an *expanding* annulus. Each is foliated by equidistant lines. In the flat annulus case, the lines are geodesics of the quadratic differential and have 0 curvature. In the latter case they have negative curvature. Rafi measures the path traveled in \mathbb{H}^2 defined by the length and twist coordinates by computing the modulus of these annuli. He shows that the distance traveled in \mathbb{H}^2 due to the expanding annulus is much smaller than the Teichmüller distance and if the reverse triangle inequality fails, it is due to the presence of an expanding annulus whose modulus is much bigger than the modulus of the flat annulus. The fact that path length in \mathbb{H}^2 is much smaller than Teichmüller length forces, by his distance formula, the presence of the domains W_i as in the statement of the theorem.

Going forward, we fix once and for all a constant M large enough to satisfy the quantitative parts of the thinness statements, the distance formula, and these reverse triangle inequality statements.

3. COMPARING MEASURES

To address averaging questions, one of course needs to consider a measure. In the present context of metric geometry, it is perhaps most natural to consider Hausdorff measure of the appropriate dimension.

Definition 14 (Hausdorff measure). The n -dimensional Hausdorff measure on a metric space will be denoted by η . It is defined by

$$\eta(E) := \lim_{\delta \rightarrow 0} \left[\inf \sum \text{diam}(U_i)^n \right],$$

where the infimum is over countable covers $\{U_i\}$ of E with $\text{diam } U_i < \delta \forall i$.

For the Teichmüller metric, there is a nontrivial h -dimensional Hausdorff measure (recalling $h = 6g - 6$). As we shall see, in order to understand average distances with respect to this measure, it will be necessary to compare with other measures, defined below, which are also natural to consider in their own right.

3.1. Measures on Finsler manifolds. The Teichmüller space carries several natural volume forms coming from its structure as a Finsler manifold. Let us discuss these general constructions first before returning to the case of $M = \mathcal{T}(S)$. The treatment closely follows the survey by Álvarez and Thompson [1].

Recall that a Finsler metric on an n -dimensional Finsler manifold M is a continuous function $F: T(M) \rightarrow \mathbb{R}$ that restricts to a norm on each tangent space $T_x(M)$. There is a dual norm on each cotangent space $T_x^*(M)$. For a point $x \in M$, let $B_x \subset T_x(M)$ and $B_x^* \subset T_x^*(M)$ denote the unit balls for these two norms. A local coordinate system (x_1, \dots, x_n) on M induces a pair of isomorphisms

$$(5) \quad \phi: T_x(M) \rightarrow \mathbb{R}^n \quad \text{and} \quad \psi: T_x^*(M) \rightarrow \mathbb{R}^n$$

defined by writing vectors and covectors with respect to the dual bases $\{\partial_{x_1}, \dots, \partial_{x_n}\}$ and $\{dx_1, \dots, dx_n\}$. By definition of the dual norm, the pairing $T_x(M) \times T_x^*(M) \rightarrow \mathbb{R}$ is sent to the standard inner product on \mathbb{R}^n under these isomorphisms. In the local coordinate chart we may now define two functions

$$f(x) = \frac{\varepsilon_n}{\lambda(\phi(B_x))} \quad \text{and} \quad g(x) = \frac{\lambda(\psi(B_x^*))}{\varepsilon_n},$$

where λ is Lebesgue measure and $\varepsilon_n := \lambda(\text{Ball}^n)$ is the Lebesgue measure of the standard unit ball in \mathbb{R}^n . While these functions clearly depend on the choice of coordinates (x_1, \dots, x_n) , one may easily check that the n -forms

$$f(x) dx_1 \wedge \dots \wedge dx_n \quad \text{and} \quad g(x) dx_1 \wedge \dots \wedge dx_n$$

are independent of the coordinate system and therefore define global volume forms on M . The former is called the *Busemann volume* on the Finsler manifold and the latter is the *Holmes–Thompson volume*; see [1] for more details. These both define measures on M .

A third measure to consider is the one induced by the canonical symplectic form ω on the cotangent bundle, defined as follows. Consider local coordinates (x_1, \dots, x_n) defined in a neighborhood $U \subset M$. The 1-forms dx_1, \dots, dx_n then give a trivialization of $T^*(M)$ over U , and we have a local coordinate system on $T^*(M)$ given by

$$(6) \quad (x_1, y_1, \dots, x_n, y_n) \mapsto \left((x_1, \dots, x_n), \sum_{i=1}^n y_i dx_i \right).$$

In these coordinates the canonical symplectic form may be written simply as $\omega = \sum dx_i \wedge dy_i$. Taking exterior powers then yields a volume form $\mu_{\text{sp}} = \omega^n/n!$ on $T^*(M)$. By restricting to the unit disk bundle $T^{*,\leq 1}(M)$ and pushing forward by the projection $\pi : T^*(M) \rightarrow M$, we obtain a *symplectic measure* \mathbf{n} on M .

Finally, a Finsler metric on a smooth manifold M^n induces a path metric d in the usual way, and this in turn gives rise to a Hausdorff measure in any dimension.

Recall that a centrally symmetric convex body $\Omega \subset \mathbb{R}^n$ determines a *polar body* $\Omega^\circ \subset (\mathbb{R}^n)^* = \mathbb{R}^n$ via

$$\Omega^\circ := \{\xi \in \mathbb{R}^n \mid \xi \cdot v \leq 1 \ \forall v \in \Omega\}.$$

The *Mahler volume* of Ω is then defined to be the product $M(\Omega) := \lambda(\Omega) \cdot \lambda(\Omega^\circ)$ of the Lebesgue volumes of Ω and Ω° . For any centrally symmetric convex body Ω , it is known that

$$(7) \quad \frac{\varepsilon_n^2}{n^{n/2}} \leq M(\Omega) \leq \varepsilon_n^2 = M(\text{Ball}^n).$$

The first inequality was established by John [9], and the latter, which gives an equality if and only if the norm is Euclidean, is known as the Blaschke–Santaló inequality [4].

Theorem 15 (Assembling facts on Finsler measures). *Suppose that M^n is a continuous Finsler manifold. Then*

- the Busemann measure μ_{B} and the n -dimensional Hausdorff measure η are equal;
- the Holmes–Thompson measure μ_{HT} and the symplectic measure \mathbf{n} are scalar multiples: $\mu_{\text{HT}} = \frac{1}{\varepsilon_n} \mathbf{n}$;
- $\mu_{\text{HT}} \leq \mu_{\text{B}} \leq (n^{n/2}) \mu_{\text{HT}}$, with equality of measures if and only if the metric is Riemannian.

Note that it is still possible for μ_{HT} and μ_{B} to be scalar multiples of each other in the non-Riemannian case, for instance on a vector space with a Finsler norm.

Proof. The first statement was originally shown by Busemann in the 1940s in [5] and is stated in modern language in [1, Thm 3.23].

The second statement is straightforward and we include a proof for completeness. Working in the local coordinates and applying the Fubini theorem, we see that the Holmes–Thompson volume of a subset $E \subset M$ is given by:

$$\begin{aligned} \int_E g(x) dx_1 \wedge \cdots \wedge dx_n &= \int_E \left(\int_{\psi(B_x^*)} \frac{1}{\varepsilon_n} d\lambda \right) dx_1 \wedge \cdots \wedge dx_n \\ &= \frac{1}{\varepsilon_n} \int_{\pi^{-1}(E) \cap T^{*,\leq 1}(M)} dy_1 \wedge \cdots \wedge dy_n \wedge dx_1 \wedge \cdots \wedge dx_n \\ &= \frac{1}{\varepsilon_n} \mathbf{n}(E). \end{aligned}$$

For the third statement, recall that the measures are defined by

$$\mu_{\text{B}}(E) = \int_E f(x) dx_1 \wedge \cdots \wedge dx_n \quad \text{and} \quad \mu_{\text{HT}}(E) = \int_E g(x) dx_1 \wedge \cdots \wedge dx_n.$$

For each $x \in M$, the unit ball $B_x \subset T_x(M)$ is sent to a centrally symmetric convex body $\phi(B_x) \subset \mathbb{R}^n$ under the isomorphism ϕ defined in (5). The polar body

is exactly given by $\phi(B_x)^\circ = \psi(B_x^*)$. Therefore, the Mahler volume of $\phi(B_x)$ is

$$M(\phi(B_x)) = \lambda(\phi(B_x)) \cdot \lambda(\psi(B_x^*)) = \varepsilon_n^2 \frac{g(x)}{f(x)}.$$

Combining with (7) now implies that $n^{-n/2}f(x) \leq g(x) \leq f(x)$ for all $x \in M$. We conclude that $\mu_{\text{HT}}(E) \leq \mu_{\text{B}}(E) \leq n^{n/2}\mu_{\text{HT}}(E)$ for all $E \subset M$. Finally, since Blaschke–Santaló can only give equality for a Euclidean norm, it follows that μ_{B} and μ_{HT} can only be equal for a Riemannian metric. \square

3.2. Measures coming from quadratic differentials. Recall that quadratic differential space $\mathcal{Q}(S)$ is naturally identified with the cotangent bundle $T^*(\mathcal{T}(S))$ of Teichmüller space, and that each quadratic differential $q \in \mathcal{Q}(S)$ has a norm $\|q\|$ given by the area of the flat structure on S induced by q . The unit disk bundle for this norm will be denoted by

$$\mathcal{Q}^{\leq 1}(S) = \{q \in \mathcal{Q}(S) : \|q\| \leq 1\}.$$

Using this disk bundle, the natural symplectic measure μ_{sp} on $\mathcal{Q}(S)$ descends to a measure \mathbf{n} on $\mathcal{T}(S)$ exactly as above. We note that ω and therefore μ_{sp} and \mathbf{n} are invariant under the action of the mapping class group.

The space $\mathcal{Q}(S)$ also carries a natural $\text{Mod}(S)$ -invariant measure μ_{hol} that is defined in terms of holonomy coordinates and which we will refer to as *holonomy measure*; it is also sometimes called Masur–Veech measure in the literature (see [12] for details). This measure has been studied extensively, for instance to establish ergodicity results for the geodesic flow. The measure μ_{hol} is also related to the “Thurston measure” μ_{TH} on the space of measured foliations \mathcal{MF} induced by the piecewise-linear structure of \mathcal{MF} [8]. Indeed, as seen in [12], μ_{hol} is equal to the pullback of $\mu_{\text{TH}} \times \mu_{\text{TH}}$ under the $\text{Mod}(S)$ -invariant map $\mathcal{Q}(S) \rightarrow \mathcal{MF} \times \mathcal{MF}$ that sends a quadratic differential to its vertical and horizontal foliations.

Just as μ_{sp} induces \mathbf{n} , the holonomy measure μ_{hol} descends to a measure \mathbf{m} on $\mathcal{T}(S)$. Explicitly, the \mathbf{m} -measure of a set $E \subset \mathcal{T}(S)$ is given by

$$\mathbf{m}(E) := \mu_{\text{hol}}(\pi^{-1}(E) \cap \mathcal{Q}^{\leq 1}(S)).$$

This measure \mathbf{m} has been studied previously in [3] and [7].

Proposition 16. [14, p.3746] *There is a scalar $k > 0$ such that $\mu_{\text{sp}} = k \cdot \mu_{\text{hol}}$.*

We recall the outlines of the argument here. In [14], it was shown that the Teichmüller flow on $\mathcal{Q}(S)$ is a Hamiltonian flow for the function

$$H(q) = \frac{\|q\|^2}{2}.$$

As such, the Teichmüller flow preserves the symplectic form ω and the corresponding measure μ_{sp} . The measures μ_{sp} and μ_{hol} both descend to the quotient space $\mathcal{Q}(S)/\text{Mod}(S)$; furthermore, the latter defines an ergodic measure for the Teichmüller flow on $\mathcal{Q}(S)/\text{Mod}(S)$ [12]. Since μ_{sp} is absolutely continuous with respect to μ_{hol} , the result follows.

We therefore also have $\mathbf{n} = k\mathbf{m}$, and combining Proposition 16 with Theorem 15 we get:

Corollary 17. *There are scalars $k_2 > k_1 > 0$ such that*

$$k_1\mathbf{m} \leq \eta \leq k_2\mathbf{m}.$$

3.3. Visual measures. The unit sphere subbundle of $\mathcal{Q}(S)$ will be denoted by

$$\mathcal{Q}^1(S) = \{q \in \mathcal{Q}(S) : \|q\| = 1\}.$$

For each $x \in \mathcal{T}(S)$, the fiber $\mathcal{Q}^1(x)$ is identified with the “space of directions” at x , and the Teichmüller geodesic flow $\varphi_t : \mathcal{Q}(S) \rightarrow \mathcal{Q}(S)$ gives rise to a homeomorphism

$$\begin{aligned} \Psi_x : \quad \mathcal{Q}^1(x) \times (0, \infty) &\rightarrow \mathcal{T}(S) \setminus \{x\} \\ (q, r) &\mapsto \pi(\varphi_r(q)), \end{aligned}$$

which serves as “polar coordinates” centered at x . Furthermore, this conjugates φ_t to a *radial flow* based at x given by

$$\hat{\varphi}_t(\pi(\varphi_r(q))) := \pi(\varphi_{r+t}(q)).$$

We will consider measures on $\mathcal{T}(S)$ that are compatible with these polar coordinates and with the radial flow.

Definition 18 (Visual measure). Given any measure κ_x on the unit sphere $\mathcal{Q}^1(x) \cong S^{h-1}$, we define the corresponding *visual measures* on $\mathcal{S}_r(x)$ and $\mathcal{T}(S)$ as follows. Firstly, the visual measure $\text{Vis}_r(\kappa_x)$ on the sphere $\mathcal{S}_r(x)$ of radius r is just the push-forward of $e^{hr}\kappa_x$ under the homeomorphism $\mathcal{Q}^1(x) \times \{r\} \cong \mathcal{S}_r(x)$. Integrating these over $(0, \infty)$ then gives a visual measure on $\mathcal{T}(S)$ defined by

$$\text{Vis}(\kappa_x)(E) := \int_{(q,r) \in E \subset \mathcal{Q}^1(S) \times (0, \infty)} e^{hr} d\kappa_x(q) d\lambda(r).$$

Said differently, $\text{Vis}(\kappa_x)$ is equal to the push-forward of $\kappa_x \times \lambda_0$ under the homeomorphism Ψ_x , where λ_0 is the weighted Lebesgue measure on $(0, \infty)$ given by $\lambda_0([a, b]) = \int_a^b e^{hr} d\lambda(r) = (e^{hb} - e^{ha})/h$. (We have scaled things in this way so that the visual measure of the ball of radius R grows like e^{hR} .)

The essential feature of visual measures is that they enjoy the following “normalized invariance” under the radial flow: For any $t \geq 0$ and measurable $E \subset \mathcal{S}_r(x)$ we have

$$\frac{\text{Vis}_{r+t}(\kappa_x)(\hat{\varphi}_t(E))}{\text{Vis}_{r+t}(\kappa_x)(\mathcal{S}_{r+t}(x))} = \frac{\text{Vis}_r(\kappa_x)(E)}{\text{Vis}_r(\kappa_x)(\mathcal{S}_r(x))}.$$

The same invariance holds for $\text{Vis}(\kappa_x)$ when we normalize with respect to annular shells $\mathcal{B}_b(x) \setminus \mathcal{B}_a(x)$ instead of spheres.

There are two visual measures that specifically interest us. Firstly, the normed vector space $\mathcal{Q}(x)$ carries a unique translation-invariant measure ν_x normalized so that $\nu_x(B_x^*) = 1$; recall that the unit ball B_x^* is just the intersection $\mathcal{Q}^{\leq 1}(S) \cap \mathcal{Q}(x)$. This induces a measure (also denoted ν_x) on the unit sphere $\mathcal{Q}^1(x)$ via the usual method of coning off: $\nu_x(E) := \nu_x([0, 1] \cdot E)$ for $E \subset \mathcal{Q}^1(x)$.

Secondly, since $\mathcal{Q}(S)$ has the structure of a fiber bundle over $\mathcal{T}(S)$, we can define a conditional measure s_x on $\mathcal{Q}(x)$ by disintegration from μ_{hol} . More precisely, s_x is the unique measure on $\mathcal{Q}(x)$ such that the μ_{hol} -measure of $E \subset \mathcal{Q}(S)$ is given by

$$\mu_{\text{hol}}(E) = \int_{\mathcal{T}(S)} s_x(E \cap \mathcal{Q}(x)) d\mathbf{m}(x).$$

Via the process of coning off, we again think of s_x as a measure on $\mathcal{Q}^1(x)$.

The space $\mathcal{Q}(S)$ of quadratic differentials is a complex vector bundle; as such, there is a natural circle action $S^1 \curvearrowright \mathcal{Q}(S)$ that preserves each fiber $\mathcal{Q}(x)$ and unit sphere $\mathcal{Q}^1(x)$. We say that a visual measure $\text{Vis}(\kappa_x)$ is *rotation-invariant* if the corresponding measure κ_x on $\mathcal{Q}^1(x)$ is invariant under this action of S^1 . The

visual measure $\text{Vis}(\nu_x)$ is rotation-invariant because S^1 preserves the unit ball B_x . Similarly, $\text{Vis}(s_x)$ is rotation-invariant because S^1 preserves μ_{hol} .

3.4. Summary. The measures on $\mathcal{T}(S)$ considered above are \mathbf{n} and \mathbf{m} (induced by the symplectic and holonomy measures on $\mathcal{Q}(S)$, respectively, via the covering map), Hausdorff measure η , the visual measures $\text{Vis}(\kappa_x)$ created by radially flowing measures on the sphere of directions $\mathcal{Q}^1(x)$, and the measures μ_{B} and μ_{HT} coming from the Finsler structure.

We found that \mathbf{n} , \mathbf{m} , and μ_{HT} are scalar multiples of each other, Hausdorff measure and Busemann measure coincide, and all five of these are mutually comparable in the sense of being bounded above and below by scalar multiples of each other. In the following section we will establish results about the structure of generic geodesic rays with respect to these measures and the visual measures.

4. VOLUME ASYMPTOTICS AND RANDOM WALKS

4.1. Volume estimates. Athreya, Bufetov, Eskin, and Mirzakhani [3] have found the following asymptotic estimate for the \mathbf{m} -volume of a ball of radius r .

Theorem 19 (Volume asymptotics [3]). *There exists a (bounded) function $f : \mathcal{T}(S) \rightarrow (0, \infty)$ such that for each $x \in \mathcal{T}(S)$*

$$\lim_{r \rightarrow \infty} \frac{\mathbf{m}(\mathcal{B}_r(x))}{e^{hr}} = f(x).$$

Corollary 20 (Definite exponential growth). *Let μ denote holonomy measure \mathbf{m} , Hausdorff measure η , or any visual measure $\mu_x = \text{Vis}(\kappa_x)$. For each $x \in \mathcal{T}(S)$, there exist constants $C_1 \leq C_2$ such that for all sufficiently large r (depending on x) we have*

$$C_1 e^{hr} \leq \mu(\mathcal{B}_r(x)) \leq C_2 e^{hr}.$$

Proof. This is built into the definition of the visual measure $\text{Vis}(\kappa_x)$. For the holonomy measure \mathbf{m} , this follows from Theorem 19 above. The same estimate then holds for Hausdorff measure η by Corollary 17. \square

For any $r > k > 0$, let $\mathcal{A}_r^k(x) = \mathcal{B}_r(x) \setminus \mathcal{B}_{r-k}(x)$ denote the annular shell between radii r and $r - k$. The fact that the volume of a ball grows exponentially in the radius means that we can focus our attention on annuli rather than on balls.

Lemma 21 (Reduction to annuli). *Fix $x \in \mathcal{T}(S)$ and let μ denote holonomy measure \mathbf{m} , Hausdorff measure η , or any visual measure $\text{Vis}(\kappa_x)$. Suppose that for all $k > 0$ we have*

$$\lim_{r \rightarrow \infty} \frac{1}{r} \frac{1}{\mu(\mathcal{A}_r^k(x))^2} \int_{\mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x)} d\mathcal{T}(y, z) d\mu(y) d\mu(z) = 2.$$

Then the same holds when $\mathcal{A}_r^k(x)$ is replaced by $\mathcal{B}_r(x)$.

Proof. Let C_1, C_2 be as in Corollary 20 above. For each k sufficiently large (satisfying $\frac{C_2}{C_1}e^{-hk} < 1$) and all sufficiently large r we have

$$\begin{aligned} 2 &\geq \frac{1}{r} \frac{1}{\mu(\mathcal{B}_r(x))^2} \int_{\mathcal{B}_r(x) \times \mathcal{B}_r(x)} d\mathcal{T}(y, z) d\mu(y) d\mu(z) \\ &\geq \left(\frac{\mu(\mathcal{B}_r(x))}{\mu(\mathcal{B}_r(x))} - \frac{\mu(\mathcal{B}_{r-k}(x))}{\mu(\mathcal{B}_r(x))} \right)^2 \frac{1}{r} \frac{1}{\mu(\mathcal{A}_r^k(x))^2} \int_{\mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x)} d\mathcal{T}(y, z) d\mu(y) d\mu(z) \\ &\geq \left(1 - \frac{C_2}{C_1} e^{-hk} \right)^2 \frac{1}{r} \frac{1}{\mu(\mathcal{A}_r^k(x))^2} \int_{\mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x)} d\mathcal{T}(y, z) d\mu(y) d\mu(z). \end{aligned}$$

The claim now follows since, by assumption, the latter becomes arbitrarily close to 2 when r and k are sufficiently large. \square

4.2. Random walks. In this section we follow the ideas of Eskin and Mirzakhani on discretizing geodesics into random paths. We begin by combining some results on the volume of balls from Athreya–Bufetov–Eskin–Mirzakhani [3] and Eskin–Mirzakhani [7].

Lemma 22 (Volume of balls [3, Theorem 1.2], [7, Lemma 3.1]). *For a fixed thick part \mathcal{T}_{ϵ_0} , there exist global constants c and C such that*

$$\mathbf{m}(\mathcal{B}_r(y)) \leq C e^{hr}$$

whenever $y \in \mathcal{T}_{\epsilon_0}$ OR $r \leq c$.

Proof. Choose a point $x \in \mathcal{T}_{\epsilon_0}$. By the volume asymptotics (Theorem 19), there exist constants R_0 and C_0 such that

$$\mathbf{m}(\mathcal{B}_r(x)) \leq C_0 e^{hr}$$

for all $r \geq R_0$. Furthermore, by increasing R_0 if necessary, we may assume that the $\text{Mod}(S)$ -translates of $\mathcal{B}_{R_0}(x)$ cover \mathcal{T}_{ϵ_0} . It follows that for any $y \in \mathcal{T}_{\epsilon_0}$ and any $r \geq 0$ we have

$$\mathbf{m}(\mathcal{B}_r(y)) \leq \mathbf{m}(\mathcal{B}_{r+R_0}(x')) = \mathbf{m}(\mathcal{B}_{r+R_0}(x)) \leq C_0 e^{hR_0} e^{hr}$$

for some $\text{Mod}(S)$ -translate x' of x . This establishes the first claim.

As for the second claim, Lemma 3.1 of [7] says that there exists constants c and C_1 such that $\mathbf{m}(\mathcal{B}_c(y)) \leq C_1$ for all $y \in \mathcal{T}(S)$. In particular, for any $r \leq c$ we have

$$\mathbf{m}(\mathcal{B}_r(y)) \leq \mathbf{m}(\mathcal{B}_c(y)) \leq C_1 \leq C_1 e^{hr}.$$

This proves the Lemma with $C = \max\{C_0 e^{hR_0}, C_1\}$. \square

To define a random walk on $\mathcal{T}(S)$ with basepoint x , first choose a net \mathcal{N} of points in $\mathcal{T}(S)$, choosing so that $x \in \mathcal{N}$, such that the points are c -separated and $(2c)$ -dense (i.e., the distances between net points are at least c but the $(2c)$ -balls about net points cover Teichmüller space).

Given a parameter τ , a *sample path of length s* (starting at x) is a map

$$\lambda: \{0, \dots, \lfloor s/\tau \rfloor\} \rightarrow \mathcal{N}$$

such that $\lambda(0) = x$ and for each index, $d\mathcal{T}(\lambda(k), \lambda(k+1)) \leq \tau$. Let $P_\tau(s)$ be the set of sample paths λ of length s , and let P_τ be the set of all sample paths of any

length $s \geq 0$. By (36) of [7], for any $\delta > 0$ and sufficiently large τ (depending on δ) we have,

$$|P_\tau(s)| \leq e^{s(h+\delta)}$$

for all $s \geq 0$. (Note that the constant C_2 in (36) of [7], coming from [7, Proposition 4.5], can be taken to equal 1.)

We now define a map $F: \mathcal{T}(S) \rightarrow P_\tau$ which takes a point y and “discretizes” the geodesic $[x, y]$ to a sample path. For any $[x, y]$ we mark off points along the geodesic starting at x and spaced by time $\tau - 2c$. For each such point we choose a nearest point in \mathcal{N} . This is the sample path associated to y . For any $\epsilon_1 > 0$, we can choose τ sufficiently large so that the image under F of the sphere $\mathcal{S}_s(x)$ is contained in $P_\tau(s(1 + \epsilon_1))$.

5. SEPARATION AND THICKNESS STATISTICS FOR GEODESIC RAYS

In order to establish distance estimates in §6, it will be necessary to restrict our attention to pairs of geodesic rays based at x that both spend a definite fraction of their time in \mathcal{T}_{ϵ_0} and stay far apart beyond some threshold radius. In this section we show that most pairs of geodesics have both of these properties, with respect to all measures we consider.

5.1. Thickness statistics. Our techniques for analyzing Teichmüller geodesics rely heavily on controlling the fraction of time they spend in the thick part \mathcal{T}_{ϵ_0} . We call this quantity the *thick-stat*; for a geodesic segment $\gamma: [0, L] \rightarrow \mathcal{T}(S)$ it is denoted by

$$\text{Th}_{\epsilon_0}^{\%}(\gamma) = \frac{|\{t \in [0, L] : \gamma(t) \in \mathcal{T}_{\epsilon_0}\}|}{L}.$$

For distinct points $x, y \in \mathcal{T}(S)$, we will use $\text{Th}_{\epsilon_0}^{\%}[x, y]$ to denote $\text{Th}_{\epsilon_0}^{\%}(\gamma)$, where $\gamma: [0, d_{\mathcal{T}}(x, y)] \rightarrow \mathcal{T}(S)$ is the Teichmüller geodesic from x to y . The goal of this subsection is to show that the thick-stat $\text{Th}_{\epsilon_0}^{\%}[x, y]$ is uniformly bounded below for most $y \in \mathcal{A}_r^k(x)$ for the measures of interest.

Definition 23. We say a measure μ on $\mathcal{T}(S)$ satisfies a *thickness estimate* if there exist constants $\delta, \epsilon_0 > 0$ such that for every $x \in \mathcal{T}(S)$, $\epsilon > 0$, and $0 < \sigma < 1$, we have

$$\mu\left(\{y \in \mathcal{A}_r^k(x) : \text{Th}_{\epsilon_0}^{\%}[x, \gamma(t)] < \delta \text{ for some } t \in [\sigma r, r]\}\right) < \epsilon$$

for sufficiently large r .

First we observe that such an estimate for visual measures is essentially immediate from the ergodicity of the Teichmüller geodesic flow φ_t . Fix a measure κ_x (either s_x or ν_x) on $\mathcal{Q}^1(x)$ and an ϵ_0 -thick part \mathcal{T}_{ϵ_0} .

Proposition 24 (Thickness statistics for visual measures). *There are constants $\delta, \epsilon_0 > 0$ such that for all $\epsilon > 0$ and all $x \in \mathcal{T}(S)$, there exist a threshold R_0 and a set $E \subset \mathcal{Q}^1(x)$ with $\kappa_x(E^c) \leq \epsilon$ so that for all $q \in E$ and $r \geq R_0$ we have $\text{Th}_{\epsilon_0}^{\%}[x, \pi(\varphi_r(q))] \geq \delta$.*

Proof. By the ergodicity of the geodesic flow [12] there is $\delta > 0$ such that the geodesic determined by almost every $q \in \mathcal{Q}(S)$ spends proportion δ of its time in \mathcal{T}_{ϵ_0} , asymptotically. The vertical foliation of each such q is uniquely ergodic [13]. If two quadratic differentials have the same vertical uniquely ergodic measured foliation then they are forwards asymptotic [11]. We conclude that almost every measured

foliation $F \in \mathcal{MF}$ (with respect to Thurston measure μ_{TH}) has the property that for any quadratic differential with vertical foliation F , the corresponding geodesic spends at least δ proportion of its time in the thick part, asymptotically.

The map $\mathcal{Q}(x) \rightarrow \mathcal{MF}$ which assigns to q its vertical foliation is a smooth map off the multiple zero locus, so it is smooth on a set of full measure. Thus it is absolutely continuous with respect to the measures κ_x and μ_{TH} . Thus the property of asymptotically spending proportion δ in the thick part is true of almost every $q \in \mathcal{Q}(x)$. Thus for every ϵ , long enough geodesics are thick for proportion δ of their length away from an exceptional set of κ_x -measure at most ϵ . \square

We next verify a thickness estimate for \mathbf{m} .

Theorem 25 (Thickness statistics for holonomy measure). *For all $0 < \theta < 1$, there exists $\epsilon_0 > 0$ such that for any $0 < \sigma < 1$, there exists $\kappa > 0$ such that for all k and all sufficiently large r ,*

$$\mathbf{m} \left(\left\{ y \in \mathcal{A}_r^k(x) : \text{Th}_{\epsilon_0}^{\%}[x, \gamma(t)] < \theta \text{ for some } t \in [\sigma r, r] \right\} \right) < C e^{hr(1-\kappa)},$$

where γ is the geodesic based at x and passing through y .

Proof. We start with a geodesic γ connecting x to $y \in \mathcal{A}_r^k(x)$. For any ϵ_1 , this is associated to a sample path λ in $\mathbf{P}_\tau(r(1+\epsilon_1))$, as in the last section. We quote the proof of Theorem 5.1 of [7] (which itself quotes [2]) to say that given θ , there exists $\delta' > 0$ and ϵ_0 such that for all large τ , with the exception of at most

$$e^{-\delta' t} e^{hr(1+\epsilon_1)}$$

sample paths λ , we have the estimate

$$\frac{1}{\lfloor t/\tau \rfloor} \left| \left\{ 1 \leq i \leq \lfloor t/\tau \rfloor : \lambda(i) \in \mathcal{T}_{\epsilon_0} \right\} \right| \geq \theta.$$

Now given σ , choose ϵ_1 small enough (forcing τ to be large) so that

$$\kappa := \frac{\delta' \sigma}{h} - \epsilon_1 > 0.$$

We conclude that for $t \geq \sigma r$ with the exception of at most

$$e^{-\delta' t} e^{hr(1+\epsilon_1)} \leq e^{hr(1-\delta' \sigma/h+\epsilon_1)} \leq e^{hr(1-\kappa)}$$

net points, the geodesics starting at x and ending at the net points have $\text{Th}_{\epsilon_0}^{\%} \geq \theta$ for ending times $t \geq \sigma r$. Now since we know that the ball of radius c has \mathbf{m} measure at most C , we are done, because all bad endpoints of geodesics $y \in \mathcal{A}_r^k(x)$ are within c of bad endpoints of sample paths. \square

This says that, except for set of endpoints y of exponentially small measure, geodesics have the property that they eventually have spent a definite fraction of their time in the thick part. This exponential bound makes the statement stronger than what is required for a thickness estimate.

5.2. Fellow-traveling statistics. We need an estimate for how likely it is for two geodesics to fellow-travel past some radius R_0 . The appropriate sort of estimate will hold provided that the measure satisfies the following property:

Definition 26. We say a measure μ on $\mathcal{T}(S)$ satisfies an *exponential decay estimate* if given any $\sigma, M_0 > 0$ there exist $C, R_0, \alpha > 0$ such that for all t, r with $R_0 \leq \sigma r \leq t \leq r$ we have

$$\frac{\mu \times \mu \left(\{(y_1, y_2) \in \mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x) : d_{\mathcal{T}}(\gamma_1(t), \gamma_2(t)) < M_0\} \right)}{\mu(\mathcal{A}_r^k(x))^2} < C e^{-\alpha t},$$

where γ_i is the geodesic ray based at x and passing through y_i .

Theorem 27 (Exponential decay for visual measures). *All rotation-invariant visual measures $\mu_x = \text{Vis}(\kappa_x)$ on $\mathcal{T}(S)$, and in particular $\text{Vis}(\nu_x)$ and $\text{Vis}(s_x)$, satisfy exponential decay estimates.*

Proof. Choose $\sigma r \leq t \leq r$, fix a point $y_1 \in \mathcal{A}_r^k(x)$, and let $E = \{y_2 \in \mathcal{A}_r^k(x) : d_{\mathcal{T}}(\gamma_1(t), \gamma_2(t)) < M_0\}$. Looking instead in the sphere $\mathcal{S}_t(x)$, we have the set $E' = \{z \in \mathcal{S}_t(x) : d_{\mathcal{T}}(\gamma_1(t), z) < M_0\}$. Notice that, by definition,

$$E = \bigcup_{s \in [r-t-k, r-t]} \hat{\varphi}_s(E').$$

(Recall that $\hat{\varphi}_s$ denotes the radial geodesic flow based at x .) Therefore, by the normalized invariance, we have

$$\begin{aligned} \mu(E) &= \int_{r-k}^r \text{Vis}_s(\kappa_x)(\hat{\varphi}_{s-t}(E')) d\lambda(s) \\ &= \int_{r-k}^r \text{Vis}_t(\kappa_x)(E') \frac{\text{Vis}_s(\kappa_x)(\mathcal{S}_s(x))}{\text{Vis}_t(\kappa_x)(\mathcal{S}_t(x))} d\lambda(s) \\ &= \frac{\kappa_x(E')}{\kappa_x(\mathcal{Q}^1(x))} \mu_x(\mathcal{A}_r^k(x)), \end{aligned}$$

where, in the last line, we have identified E' with its image in $\mathcal{Q}^1(x) \cong \mathcal{S}_t(x)$.

It remains to find C, R_0 (independent of y_1) such that $\kappa_x(E')/\kappa_x(\mathcal{Q}^1(x)) \leq C e^{-t}$ when $t \geq R_0$. Recall that S^1 acts freely on $\mathcal{Q}^1(x)$ by rotations. Choosing orbit representatives, we may realize $\mathcal{Q}^1(x)$ as a setwise product $(\mathcal{Q}^1(x)/S^1) \times S^1$. The measure κ_x pushes forward to a measure on $\mathcal{Q}^1(x)/S^1$. By disintegration, we then obtain a measure on each fiber S^1 which, by the rotation-invariance of κ_x , must agree with Lebesgue measure up to a scalar. For any two points $z, z' \in E'$, the triangle inequality gives $d_{\mathcal{T}}(z, z') \leq 2M_0$. Now suppose that z and z' lie in the same Teichmüller disk, meaning that the unit quadratic differentials associated to the geodesics $[x, z]$ and $[x, z']$ lie in the same S^1 -orbit. Each Teichmüller disk is an isometrically embedded copy of the hyperbolic plane. Thus, when t is large compared to M_0 , hyperbolic geometry implies that the fraction of each S^1 -orbit contained in E' is at most $C e^{-t}$ for some constant C . Using the product structure and integrating over the $\mathcal{Q}^1(x)/S^1$ factor, Fubini's theorem then implies that $\kappa_x(E')/\kappa_x(\mathcal{Q}^1(x)) \leq C e^{-t}$ as well. \square

Theorem 28 (Exponential decay for holonomy measure). *The measure \mathbf{m} satisfies an exponential decay estimate.*

Proof. For a given r, t , let

$$D_{r,t} = \{(y_1, y_2) \in \mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x) : d_{\mathcal{T}}(\gamma_1(t), \gamma_2(t)) < M_0\}$$

denote the the set in question. Set $\theta = 1/2$ and choose ϵ_0 and $\kappa > 0$ as in Theorem 25 so that the \mathbf{m} -measure of the set

$$E_r = \{y_1 \in \mathcal{A}_r^k(x) : \text{Th}_{\epsilon_0}^\%[x, \gamma_1(t)] < \theta \text{ for some } t \in [\sigma r, r]\}$$

is at most $Ce^{hr(1-\kappa)}$ for all large r . We may now write $D_{r,t} = D'_{r,t} \cup D''_{r,t}$, where

$$D'_{r,t} = \{(y_1, y_2) \in D_{r,t} : y_1 \in E_r\} \quad \text{and} \quad D''_{r,t} = \{(y_1, y_2) \in D_{r,t} : y_1 \notin E_r\}.$$

By the above, we have that

$$(8) \quad \frac{\mathbf{m} \times \mathbf{m}(D'_{r,t})}{\mathbf{m}(\mathcal{A}_r^k(x))} \leq Ce^{hr(1-\kappa)} \leq Ce^{hr} e^{-h\kappa t}$$

for all large r and all $t \leq r$.

Choose any point $y_1 \in \mathcal{A}_r^k(x) \setminus E_r$ and fix some $t \in [\sigma r, r]$. Then any point $y_2 \in \mathcal{A}_r^k(x)$ satisfying $d_{\mathcal{T}}(\gamma_1(t), \gamma_2(t)) < M_0$ must be contained in the ball of radius $r - t + M_0$ about $\gamma_1(t)$. Since $t \geq \sigma r$ and $y_1 \notin E_r$, we know that $\text{Th}_{\epsilon_0}^\%[x, \gamma_1(t)] \geq \theta$. Choosing any $0 < \alpha < \theta$, there must exist a time $t' \in [\alpha t, t]$ for which $\gamma_1(t') \in \mathcal{T}_{\epsilon_0}$. Furthermore, each such y_2 lies within the ball of radius $(1 - \alpha)t + M_0 + r - t$ about $\gamma_1(t')$. Applying Lemma 22, we see that

$$(9) \quad \frac{\mathbf{m} \times \mathbf{m}(D''_{r,t})}{\mathbf{m}(\mathcal{A}_r^k(x))} \leq \mathbf{m}(\mathcal{B}_{r-\alpha t+M_0}(\gamma_1(t'))) \leq Ce^{hM_0} e^{hr} e^{-h\alpha t}.$$

The claim now follows from (8), (9) and the fact that, by Corollary 20, there exists a constant C_1 such that $\mathbf{m}(\mathcal{A}_r^k(x)) \geq C_1 e^{hr}$ for all large r . \square

Finally, we see that for any measure satisfying an exponential decay estimate, most pairs of geodesic rays are in fact *never* near each other beyond some threshold.

Proposition 29 (Separation is forever). *Suppose a measure μ satisfies an exponential decay estimate. Then for any $0 < \sigma < 1$ and $M_0 > 0$ there exist $R_0, C, \alpha > 0$ such that for all $r \geq R_0$ the $\mu \times \mu$ -measure of the set*

$$\{(y_1, y_2) \in \mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x) : d_{\mathcal{T}}(\gamma_1(t), \gamma_2(t)) < M_0 \text{ for some } t \in [\sigma r, r]\}$$

is bounded by $\mu(\mathcal{A}_r^k(x))^2 C e^{-\alpha \sigma r}$, where here γ_i denotes the geodesic ray based at x and passing through y_i .

Proof. If (y_1, y_2) is such a point, then there is some $n \in \mathbb{N}$, $n \leq (1 - \sigma)r$ such that $d_{\mathcal{T}}(\gamma_1(\sigma r + n), \gamma_2(\sigma r + n)) < M_0 + 2$. Thus our set of points is contained in the union of exceptional sets corresponding to the radii $\sigma r, \sigma r + 1, \dots, \sigma r + \lfloor r - \sigma r \rfloor$. Using the exponential decay estimate, we see that our set has measure at most

$$\begin{aligned} & \mu(\mathcal{A}_r^k(x))^2 C \left(e^{-\alpha \sigma r} + e^{-\alpha(\sigma r + 1)} + \dots + e^{-\alpha(\sigma r + \lfloor r - \sigma r \rfloor)} \right) \\ & < \mu(\mathcal{A}_r^k(x))^2 C \left(\frac{1}{1 - e^{-\alpha}} \right) e^{-\alpha \sigma r}. \end{aligned} \quad \square$$

Thus we can conclude that after throwing out a subset of $\mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x)$ of measure which is an arbitrarily small proportion, all pairs of geodesics stay separated by an arbitrarily chosen distance in Teichmüller space after a threshold time σr has elapsed. Later we will show that after waiting even longer we may also assume that every pair of geodesics has big curve complex distance. This relies both on the large Teichmüller distance established above and the fact that most geodesics spend a definite fraction of their time in the thick part.

6. DISTANCE ESTIMATES

6.1. Progress in the curve complex. The goal of this subsection is to prove Proposition 38, which says that any geodesic that spends a definite fraction of its time in the thick part must move at a definite rate in the curve complex. The idea is that long subintervals contained in \mathcal{T}_{ϵ_0} contribute to progress in $\mathcal{C}(S)$; alternately, one could consider intervals in the complement of $\bigcup_V T_V$. For this analysis, we would like to bound the number of connected components of $\bigcup_V T_V$ in terms of $d_S(x, y)$. One bound is given by the number of nonempty thin intervals. While there may be arbitrarily many such T_V , some of these will be redundant in the sense that $T_V \subset T_W$ for some other subsurface W .

Definition 30 (Thin-significance). Recall the choice of global constant M . A proper subsurface $V \subsetneq S$ is said to be *thin-significant* for the geodesic segment $[x, y]$ if $d_{\mathcal{C}(V)}(x, y) \geq 3M$ and for every other proper subsurface $Z \subsetneq S$ with $d_{\mathcal{C}(Z)}(x, y) \geq 3M$ we have $T_V \not\subset T_Z$.

Remark. In this subsection we will focus on the curve complex distance $d_{\mathcal{C}(V)}$ for a subsurface V . Recall that this agrees with the usual projection distance d_V in the case that V is non-annular, but that $d_{\mathcal{C}(A)}$ and d_A differ for annuli. We will take care to handle exceptional annuli carefully.

Our next goal is to bound the number of thin-significant subsurfaces along an arbitrary geodesic. For this, we will use the work of Rafi–Schleimer [20] bounding the size of an antichain in the poset of subsurfaces of S .

Definition 31 (Antichain). Given a subsurface $\Sigma \subset S$ a pair of points $x, y \in \mathcal{T}(S)$ and constants $T_1 \geq T_0 > 0$, a collection Ω of proper subsurfaces of Σ is an *antichain* for (Σ, x, y, T_0, T_1) if the following hold:

- if $Y, Y' \in \Omega$, then Y is not a proper subsurface of Y' ;
- if $Y \in \Omega$, then $d_{\mathcal{C}(Y)}(x, y) \geq T_0$; and
- if $Z \subsetneq \Sigma$ and $d_{\mathcal{C}(Z)}(x, y) \geq T_1$, then $Z \subset Y$ for some $Y \in \Omega$.

Lemma 32 (Antichain bound [20, Lem 7.1]). *For every $\Sigma \subset S$ and sufficiently large $T_1 \geq T_0 > 0$, there is a constant $A = A(\Sigma, T_0, T_1)$ so that if Ω is an antichain for (Σ, x, y, T_0, T_1) then*

$$|\Omega| \leq A \cdot d_{\mathcal{C}(\Sigma)}(x, y).$$

We now prove a proposition showing that if there are a large enough number of thin-significant subsurfaces along a geodesic, then the image of the geodesic makes definite progress in the curve complex. The following notation will be used in the proof.

Definition 33. Consider a geodesic segment $[x, y] \subset \mathcal{T}(S)$ and a collection Ω of proper subsurfaces of S . We will consider three partial orders on the set Ω :

- (1) $V \leq_1 W \iff V \subset W$,
- (2) $V \leq_2 W \iff T_V \subset T_W$, and
- (3) $V \leq_3 W \iff V \subset W \text{ and } T_V \subset T_W$.

The subcollection of Ω consisting of maximal elements with respect to \leq_* will be denoted $(\Omega)_*$; notice that these sets are related by $(\Omega)_1 \subset (\Omega)_3 \supset (\Omega)_2$. Elements of $(\Omega)_1$ are said to be *topologically maximal* with respect to Ω .

Proposition 34 (Progress from thin-significant subsurfaces). *For any t_0 , there is a constant N such that if $d_{C(S)}(x, y) \leq t_0$, then the number of thin-significant subsurfaces Y along $[x, y]$ is at most N .*

Proof. Let $\Omega = \{V \subsetneq S : d_{C(V)}(x, y) \geq 3M\}$ be the collection of proper subsurfaces which have a large projection. By definition, the set of thin-significant subsurfaces is exactly given by $(\Omega)_2$. On the other hand, the subcollection $(\Omega)_1$ of topologically maximal subsurfaces clearly forms an antichain for $(S, x, y, 3M, 3M)$. By Lemma 32, we therefore have $|(\Omega)_1| \leq At_0$ for some constant A . We will extend this to a bound on the cardinality of the larger set $(\Omega)_3$; this will imply the proposition because $(\Omega)_2 \subset (\Omega)_3$.

Fix a proper subsurface $W \in \Omega$ and consider the set $\mathcal{U}_W = \{V \in (\Omega)_3 : V \subsetneq W\}$. We claim that $|\mathcal{U}_W|$ is bounded by a constant depending only on the complexity of W . By the above, this will suffice because each $V \in (\Omega)_3$ is either equal to or properly contained in some topologically maximal proper subsurface $W \in (\Omega)_1$.

First consider those $V \in \mathcal{U}_W$ for which $T_V \cap T_W \neq \emptyset$. The definition of \leq_3 implies that $T_V \not\subset T_W$; therefore T_V must overlap with at least one endpoint of T_W . If T_{V_1} and T_{V_2} both contain the initial endpoint of T_W , then $T_{V_1} \cap T_{V_2} \neq \emptyset$ and so we cannot have $V_1 \pitchfork V_2$. Since there is a universal bound on the number of subsurfaces such that no two intersect transversely, this bounds the number of $V \in \mathcal{U}_W$ for which $T_V \cap T_W \neq \emptyset$.

It remains to bound the number of $V \in \mathcal{U}_W$ for which $T_V \cap T_W = \emptyset$; we will only focus on the case $T_V < T_W$. Suppose that $T_W = [a, b] \subset [x, y]$ and consider the set $\Omega' = \{V \in \Omega : V \subsetneq W \text{ and } d_{C(V)}(x, a) \geq 2M\}$. Notice that the subcollection $(\Omega')_1$ forms an antichain for $(W, x, a, 2M, 4M)$: the only difficulty is to check that every $Y \subsetneq W$ with $d_{C(Y)}(x, a) \geq 4M$ is contained in an element of $(\Omega')_1$. However, this is true because the reverse triangle inequality guarantees that $d_{C(Y)}(x, y) \geq d_{C(Y)}(x, a) - B \geq 3M$ and therefore that $Y \in \Omega'$. Since $d_{C(W)}(x, a) \leq M$, Lemma 32 now gives a bound on $|(\Omega')_1|$. Finally, notice that for each $V \in \mathcal{U}_W$ with $T_V < T_W$ the triangle inequality gives $d_{C(V)}(x, a) \geq d_{C(V)}(x, y) - M \geq 2M$ and so ensures that $V \in \Omega'$. Therefore each such V is contained in some topologically maximal $Z \in \Omega'$; that is to say, each $V \in \mathcal{U}_W$ with $T_V < T_W$ is contained in \mathcal{U}_Z for some $Z \in (\Omega')_1$. The bound on $|\mathcal{U}_W|$ now follows by induction on the complexity of the subsurface W . \square

Definition 35. Define a constant $P := 36 \max \{\log_+(1/\epsilon_0), \log_+(3M)\}$. Say that an annular subsurface A has an *exceptional thin interval* T_A along $[x, y]$ if $d_{C(A)}(x, y) \leq 3M$ but $d_A(x, y) \geq P$. By Lemma 9, this is only possible if $l_x(\partial A) < \epsilon_0$ or $l_y(\partial A) < \epsilon_0$. Therefore, such an annulus must determine a nonempty thin interval along $[x, y]$ that contains either x or y . Since all annuli A with $l_x(\partial A) < \epsilon_0$ must be disjoint, we see that there is a universal bound (namely $6g - 6 = h$) on the number of annuli with exceptional thin intervals along an arbitrary geodesic $[x, y]$.

We now define the *primary thin portion* \mathcal{W} of a geodesic segment $[x, y]$ to be the union of thin intervals T_V for all non-annular proper subsurfaces with $d_V(x, y) \geq 3M$ and all annular subsurfaces with $d_{C(A)}(x, y) \geq 3M$ or $d_A(x, y) \geq P$. In the case that $d_S(x, y) \leq t_0$, Proposition 34 implies that \mathcal{W} is the union of at most $N + h$ thin intervals, namely, those corresponding to the thin-significant subsurfaces and to annuli with exceptional thin intervals.

While \mathcal{W} does contain most of the nonempty thin intervals along the geodesic, it need not cover the entire time that $[x, y]$ spends in the thin part of Teichmüller space. Nevertheless, the projections to all proper subsurfaces remain uniformly bounded on the complement of \mathcal{W} .

Lemma 36 (Complement of \mathcal{W}). *There exists a constant M' with the following property. If $[a, b] \subset [x, y] \setminus \mathcal{W}$ is a connected interval in the complement of the primary thin portion of $[x, y]$, then $d_Y(a, b) \leq M'$ for all proper subsurfaces $Y \subsetneq S$.*

Proof. First suppose that Y satisfies the reverse triangle inequality (4) along $[x, y]$. If Y is non-annular and $d_Y(x, y) \geq 3M$, or if Y is an annulus and $d_{C(Y)}(x, y) \geq 3M$ or $d_Y(x, y) \geq P$, then $T_Y \subset \mathcal{W}$ by definition. Therefore $d_Y(a, b) \leq M$ since $[a, b] \cap T_Y = \emptyset$. If this is not the case, then the reverse triangle inequality gives $d_Y(a, b) \leq d_Y(x, y) + B \leq 4M + P$ as claimed.

It remains to consider an annular subsurface $A \subset S$ for which the reverse triangle inequality fails. We may assume that $d_{C(A)}(x, y) \leq 3M$ and $d_A(x, y) \leq P$, for otherwise we have $T_A \subset \mathcal{W}$ and $d_A(a, b) \leq M$ as above. Let B' be the constant corresponding to the threshold $5M + P$ in Theorem 13 (R.T.I. exception). According to that theorem, applied to the geodesic $[a, y]$, we either have $d_A(a, b) + d_A(b, y) \leq d_A(a, y) + B'$, or there exist subsurfaces W_i that satisfy the reverse triangle inequality and which have $d_{W_i}(a, b) \geq 5M + P$. However, as we have seen above, there are no such proper subsurfaces. Therefore the former inequality must hold. We similarly have $d(x, a) + d(a, b) \leq d(x, b) + B'$. Adding these inequalities and using the triangle inequality then gives $d_A(a, b) \leq d_A(x, y) + B' \leq P + B'$. \square

By the distance formula (1), it follows that long intervals in the complement of \mathcal{W} must travel a large distance the curve complex $\mathcal{C}(S)$ of the whole surface. The following lemma says that each such subinterval contributes to the curve complex distance along the total geodesic.

Lemma 37 (Cumulative contribution of subintervals). *There exist constants $0 < \rho_1 < 1$ and $D_1 > 0$ such that for all $d > D_1$, if $[x, y]$ is a Teichmüller geodesic that contains n subintervals $[x_i, y_i]$ with disjoint interiors whose endpoints satisfy $d_S(x_i, y_i) \geq d$, then*

$$d_S(x, y) \geq \rho_1 n d.$$

Proof. Applying the reverse triangle inequality (4) to the points x_i and y_i we have $d_S(x, x_i) + d_S(x_i, y_i) + d_S(y_i, y) \leq d_S(x, y) + 2B$. By recursively applying this observation to $[x, x_i]$ and $[y_i, y]$ and then throwing out the complementary intervals, we find have that

$$d_S(x, y) \geq \sum d_S(x_i, y_i) - 2nB \geq nd - 2nB.$$

Choose $D_1 > 4B$ and $\rho_1 = 1/2$. Then for $d \geq D_1$ the quantity on the right side is at least $\rho_1 nd$. \square

We now fix once and for all a “definite progress” constant $D > 0$ sufficiently large so that $\rho_1 D > D_1$ (and thus $D > D_1$ as well). Applying the distance formula (1) with the threshold M' given by Lemma 36, we have quasi-isometry constants K, C such that $d_{\mathcal{T}}(a, b) \leq K d_S(a, b) + C$ for any connected interval $[a, b] \subset [x, y] \setminus \mathcal{W}$. This gives rise to a fixed value L such that any interval $[a, b]$ of length at least L that lies entirely in $[x, y] \setminus \mathcal{W}$ satisfies $d_S(a, b) \geq D$; for example, any value $L \geq KD + C$ will suffice. Thus according to Lemma 37, if I is any interval along a geodesic that

contains a subinterval of length L that is disjoint from \mathcal{W} , then the distance in the curve complex between the endpoints of I is at least $\rho_1 D$.

Furthermore by Proposition 34 associated to the constant $t_0 = \rho_1 D$, there is a constant N so that the conclusion of Proposition 34 holds.

Proposition 38 (Definite progress). *For each $0 < \delta < 1$, there exist constants $\rho, R_1 > 0$ with the following property. If $[x, y]$ is a Teichmüller geodesic of length $r \geq R_1$ such that $\text{Th}_{\epsilon_0}^\%[x, y] \geq \delta$, then $d_S(x, y) \geq \rho r$.*

Proof. Let N denote the constant obtained by applying Proposition 34 with $t_0 = \rho_1 D$. Choose n so that $n\delta > 1$ and make the following definitions:

$$\delta' = \frac{n\delta - 1}{n - 1}, \quad T_0 \geq \frac{L(N + h + 1)}{\delta'}, \quad R_1 = 2T_0, \quad \rho = \frac{\rho_1^2 D}{2nT_0}.$$

Let $[x, y]$ be a Teichmüller geodesic of length $r \geq R_1$ that spends at least δr in the thick part. Set $m = \lfloor r/T_0 \rfloor$ and divide $[x, y]$ into m subsegments of length $r/m \geq T_0$. Let us say that a subsegment $[a, b] \subset [x, y]$ is *stalled* if $d_S(a, b) < \rho_1 D$ and *progressing* if $d_S(a, b) \geq \rho_1 D$. Suppose that m_1 of the subsegments are stalled, and thus $m_2 = m - m_1$ are progressing. Given a stalled segment $[a, b]$, we decompose it into its primary thin portion \mathcal{W} and the corresponding complementary subintervals. Since the interval is stalled, Proposition 34 ensures that \mathcal{W} is the union at most $N + h$ thin intervals. Therefore we can conclude that \mathcal{W} has at most $N + h + 1$ complementary subintervals in $[a, b]$. Furthermore, each complementary subinterval has length at most L , for otherwise we would have $d_S(a, b) \geq \rho_1 D$ by the preceding paragraph. Since \mathcal{W} is contained in the thin part, we see that the total amount of time that this interval $[a, b]$ spends in the thick part is at most

$$(N + h + 1)L \leq \delta' T_0 \leq \delta' r/m.$$

Therefore the total amount of time that the full interval $[x, y]$ spends in the thick part is at most

$$\left(\delta' \frac{r}{m}\right) m_1 + \left(\frac{r}{m}\right) m_2 = \frac{r}{m} (\delta' m_1 + m_2).$$

We claim that $m_2 \geq m/n$. If this were not the case, then we necessarily have $m_1 > (n - 1)m/n$. Since $\delta' < 1$, it follows that

$$\delta' \cdot m_1 + 1 \cdot m_2 < \delta' \cdot m \frac{n - 1}{n} + 1 \cdot m \frac{1}{n}$$

where the inequality is valid by the elementary fact that for any constants $a, b, c, d, \alpha, \beta$ such that $a + b = c + d$ and $0 < \alpha < \beta$ we have

$$(10) \quad \alpha \cdot a + \beta \cdot b < \alpha \cdot c + \beta \cdot d \iff a > c.$$

But then the amount of time that $[x, y]$ is thick is less than

$$\frac{r}{m} \left(\delta' m \frac{n - 1}{n} + m \frac{1}{n} \right) = r \left(\frac{n\delta - 1}{n - 1} \cdot \frac{n - 1}{n} + \frac{1}{n} \right) = r\delta,$$

which contradicts the assumption on $[x, y]$. Therefore $m_2 \geq m/n$, as claimed.

On each of the m_2 progressing intervals, the curve complex distance between endpoints is at least $\rho_1 D$. Therefore, cumulative contribution of subintervals (Lemma 37) implies that

$$d_S(x, y) \geq \rho_1 m_2 (\rho_1 D) \geq \rho_1^2 D \frac{m}{n} \geq \frac{\rho_1^2 D}{n} \left(\frac{r}{T_0} - 1 \right) \geq \frac{\rho_1^2 D}{2nT_0} r = \rho r. \quad \square$$

6.2. Separation in the curve complex. We next establish a technical lemma which says that if we have a pair of geodesic segments from a common basepoint, we can “back up” from the endpoints to earlier points that admit distance estimates for both $d_{\mathcal{T}}$ and d_S .

Lemma 39 (Backing up to eliminate thin parts). *There exist constants k_0 and M_0 with the following property. Suppose γ_1, γ_2 are a pair of geodesic rays based at x and let $y = \gamma_1(t_1)$ and $z = \gamma_2(t_2)$. Then there are times $0 \leq t'_i \leq t_i$ and corresponding points $y' = \gamma_1(t'_1)$, $z' = \gamma_2(t'_2)$ such that*

- (i) $d_S(y, y') \leq k_0 \cdot d_S(y, z)$ and $d_S(z, z') \leq k_0 \cdot d_S(y, z)$;
- (ii) either $d_S(y', z') \leq 6$ or $d_Y(y', z') \leq M_0$ for all proper subsurfaces Y .

We will call these the backup points and backup times for the segments $[x, y], [x, z]$. Furthermore,

- (iii) given any d , there are N_0 and c_0 such that if $d_{\mathcal{T}}(y, z) \geq N_0$ and $d_S(y, z) \leq d$, then the backup points satisfy

$$d_{\mathcal{T}}(y, y') \geq c_0 \cdot d_{\mathcal{T}}(y, z) \quad \text{or} \quad d_{\mathcal{T}}(z, z') \geq c_0 \cdot d_{\mathcal{T}}(y, z)$$

Conclusion (i) says that the backup points are not much farther from the endpoints in the curve complex than the endpoints are from each other. The interpretation of (ii) is that in the distance formula (1), a significant contribution is made either by the whole curve complex distance or by projection distances to proper subsurface (but not both). These properties of backup points hold in general; (iii) says that if the Teichmüller distance between endpoints is long enough relative to their curve-complex distance, then on at least one side the distance backed up was significant. (Compare (i) and (iii).)

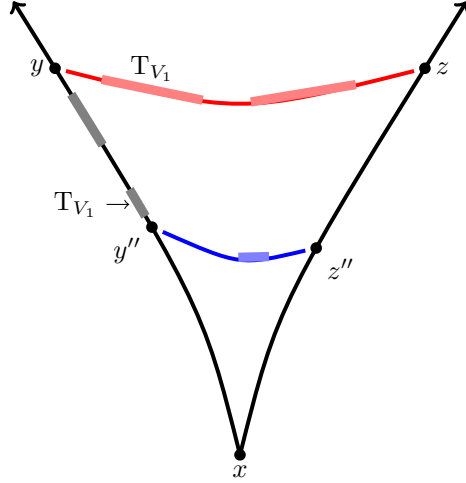


FIGURE 2. Two rays. Some thin intervals are shown, and two thin intervals for the same subsurface $V_1 \subset \Omega$ are marked.

Proof. First, we will construct backup points satisfying (i) and (ii). Then we will verify (iii).

Let

$$\Omega = \{V \subsetneq S : d_V(y, z) \geq 10M\}.$$

If $\Omega = \emptyset$, let $y' = y$ and $z' = z$ and we are done, since (i) is trivially satisfied and (ii) works with $M_0 = 10M$. So assume Ω is nonempty.

For each $V \in \Omega$, there is a thin interval T_V along $[y, z]$. By the triangle inequality, for each $V \in \Omega$, either $d_V(x, y)$ or $d_V(x, z)$ is at least $5M$ and so there is a nonempty thin interval T_V along at least one of $[x, y]$ or $[x, z]$, as in Figure 2.

Now let y'' and z'' be the earliest endpoints on γ_i of any T_V for $V \in \Omega$. That is, we move back in time until we have passed through each thin interval T_V for $V \in \Omega$. Let us use the notation V_1 and V_2 for the last surfaces in Ω whose thin interval one passes through in going from y to y'' and z to z'' , respectively. Thus we have a short curve at y'' which is also short somewhere along $[y, z]$, which means that the image of y'' under the projection $\mathcal{T}(S) \rightarrow \mathcal{C}(S)$, which sends a surface to its Bers marking, shares a point with the image of $[y, z]$ in the curve complex. By the reverse triangle inequality, this means $d_S(y, y'')$ is bounded above relative to $d_S(y, z)$, as in (i).

Now with respect to the new points y'' and z'' , let

$$\Omega'' = \{Y : d_Y(y'', z'') \geq 100M + 2B'\},$$

where B' is the constant obtained by applying Theorem 13 (R.T.I. exception) with the threshold $50M$. Note that $\Omega \cap \Omega'' = \emptyset$ because no $V \in \Omega$ has a thin interval along $[x, y'']$ or $[x, z'']$, so we would get a violation of the triangle inequality if V were also in Ω'' . If $\Omega'' = \emptyset$, then by setting $y' = y''$, $z' = z''$, and letting $M_0 = 100M + 2B'$, we are again done. So assume $\Omega'' \neq \emptyset$.

Fix some $Y \in \Omega''$. We know that either $d_Y(x, y'')$ or $d_Y(x, z'')$ is at least $50M + B'$; let us suppose that the first of these is true. First assume that Y satisfies the reverse triangle inequality (Lemma 12). Then

$$d_Y(x, y) \geq 50M + B' - B \geq 49M.$$

Since $d_Y(y, z) < 10M$, the triangle inequality then implies

$$d_Y(x, z) \geq d_Y(x, y) - 10M \geq 39M.$$

That is, for any $Y \in \Omega''$ that satisfies the reverse triangle inequality we have $d_Y(x, y), d_Y(x, z) \geq 39M$.

If Y is an exception to the reverse triangle inequality, it must be an annulus, say with core curve α . Then by Theorem 13 we either have $d_Y(x, y) \geq d_Y(x, y'') - B'$, or there exists W disjoint from α such that W satisfies the reverse triangle inequality along $[x, y]$ and $d_W(x, y'') \geq 50M$. In the former case we have $d_Y(x, y) \geq 50M$ so that $d_Y(x, z) \geq 40M$ by the triangle inequality. In the latter case, we have $d_W(x, y) \geq d_W(x, y'') - B \geq 49M$ by (4). Furthermore, since W must have a thin interval along $[x, y'']$, it cannot be in Ω and so we conclude $d_W(x, z) \geq 39$ as above.

In summary, to each $Y \in \Omega''$ we have associated a domain $W \notin \Omega$ (equal to Y except in the one case) such that both $d_W(x, y) \geq 39M$ and $d_W(x, z) \geq 39M$. Notice that we need only resort to the exceptional case $W \neq Y$ when $d_Y(y'', y) \geq M$ (still assuming here that $d_Y(x, y'') \geq 50M + B'$), for otherwise we may again conclude $d_Y(x, z) \geq 39$ by the triangle inequality. The inequalities $d_Y(x, y''), d_Y(y'', y) \geq M$ require that Y is thin along both $[x, y'']$ and $[y'', y]$; by the connectedness of T_Y , this implies that $\alpha = \partial Y$ is thin at y'' and consequently disjoint from ∂V_1 . Therefore,

since W is disjoint from α , we may safely assume that $d_S(\partial W, \partial V_1) \leq 2$ in the exceptional case $W \neq Y$.

We now make the following observation:

$$(11) \quad Z \notin \Omega, Z \cap V_1 \text{ and } d_Z(x, y) \geq 39M \implies T_Z < T_{V_1} \text{ along } \gamma_1,$$

and likewise for z, V_2, γ_2 . We argue by contradiction. Otherwise $T_{V_1} < T_Z$ and so we have $d_Z(x, y'') \leq M$ which implies $d_Z(y'', y) \geq 38M$. Since $d_Z(y, z) < 10M$ and $d_{V_1}(y, z) \geq 10M$ this would violate Lemma 11 viewed from y . Namely along $[y, y'']$ we pass through T_Z and then T_{V_1} while along $[y, z]$ we pass through T_Z with a much smaller projection.

For each $Y \in \Omega''$, we now see that the associated subsurface $W \notin \Omega$ satisfies $d_S(\partial W, \partial V_i) \leq 2$ for either $i = 1$ or $i = 2$. We have already observed this when $W \neq Y$, and in the case that $W = Y$ we in fact have that Y is disjoint or nested with respect to either V_1 or V_2 . Indeed, if $Y \cap V_i$ for both $i = 1, 2$, then (11) would imply that $T_Y \subset [x, y'']$ and $T_Y \subset [x, z'']$. But then $d_Y(y'', y), d_Y(z'', z) \leq M$ and since $d_Y(y, z) \leq 10M$ we have contradicted $d_Y(y'', z'') \geq 50M$.

We are ready to define the back up points y' and z' . Choose any $Y \in \Omega''$ (recall that we are assuming $\Omega'' \neq \emptyset$) and let W be the associated subsurface. If W is disjoint or nested with respect to V_1 then define $y' = y''$. Otherwise (11) implies that $T_W < T_{V_1}$ along γ_1 ; in this case we back up farther and define y' to be the beginning of T_W along $[x, y]$. Define the point z' similarly.

Since either V_1 or W is thin at y' and $d_S(\partial W, Z) \leq 2$ for some surface that is thin along $[y, z]$ (namely, V_1 or V_2), we conclude that $d_S(y, y')$ (and similarly $d_S(z, z')$) is bounded relative to $d_S(y, z)$, as in (i). If we have backed up farther on both sides, then W is thin at both y' and z' so that $d_S(y', z') \leq 4$ (recall that a Bers marking has diameter 2 in $\mathcal{C}(S)$). If we backed up on just one side, then we have a path of length 1 in the curve complex (either $\partial W - \partial V_2$ or $\partial V_1 - \partial W$) showing that $d_S(y', z') \leq 5$. Finally, if we didn't back up on either side, then the path $\partial V_1 - \partial W - \partial V_2$ of length 2 in the curve complex implies $d_S(y', z') \leq 6$. This verifies (ii).

We now consider (iii), so we are assuming that $d_S(y, z) \leq d$. The distance formula (1) says

$$d_{\mathcal{T}}(y, z) \leq K \left(d_S(y, z) + \sum_{\Omega} d_V(y, z) \right) + C \leq Kd + C + K \sum_{\Omega} d_V(y, z)$$

where K, C are the constants coming from threshold $M_0 = 10M$. Let us write d_{12} for $d_{\mathcal{T}}(y, z)$. Take $N_0 \geq 2Kd + 2C$, so that the assumption of (iii) says that $d_{12} \geq 2Kd + 2C$. This gives

$$\sum_{\Omega} d_V(y, z) \geq \frac{d_{12} - Kd - C}{K} \geq \frac{d_{12} - d_{12}/2}{K} = \frac{d_{12}}{2K}.$$

For each $V \in \Omega$, the definition of y' and z' ensures that $T_V \cap [x, y'] = T_V \cap [x, z'] = \emptyset$; therefore $d_V(y', z') \leq 2M$. Therefore

$$d_V(y, z) \leq d_V(y, y') + d_V(z, z') + 2M.$$

Furthermore, since $d_V(y, z) \geq 10M$, it follows that

$$d_V(y, y') + d_V(z, z') \geq d_V(y, z) - 2M \geq \frac{4}{5}d_V(y, z).$$

Notice that it cannot be the case that $d_V(y, y') \leq 3M$ and $d_V(z, z') \leq 3M$, for this would imply that $d_V(y, z) \leq 3M + 3M + 2M < 10M$, which is not the case. Therefore at least one of $d_V(y, y')$ or $d_V(z, z')$ is larger than $3M$. We now have that

$$\begin{aligned} [d_V(y, y')]_{3M} + [d_V(z, z')]_{3M} &\geq d_V(y, y') + d_V(z, z') - 3M \\ &\geq \frac{4}{5}d_V(y, z) - 3M \geq \left(\frac{4}{5} - \frac{3}{10}\right)d_V(y, z) = \frac{1}{2}d_V(y, z). \end{aligned}$$

Let K', C' be the constants in (1) for the threshold $M_0 = 3M$ and enlarge N_0 if necessary to ensure that $d_{12} \geq 16C'K'K$. The distance formula then gives

$$\begin{aligned} d_{\mathcal{T}}(y, y') + d_{\mathcal{T}}(z, z') &\geq \frac{1}{K'} \left(d_S(y, y') + d_S(z, z') + \sum_{\Omega} [d_V(y, y')]_{3M} + [d_V(z, z')]_{3M} \right) - 2C' \\ &\geq \frac{1}{K'} \sum \frac{1}{2}d_V(y, z) - 2C' \geq \frac{d_{12}}{4K'K} - 2C' \\ &\geq \frac{d_{12}}{4K'K} - \frac{d_{12}}{8K'K} = \frac{d_{12}}{8K'K}, \end{aligned}$$

so we are done if we take $c_0 = \frac{1}{16K'K}$. \square

Next we show that a pair of long geodesic segments which both stay far apart from each other in $\mathcal{T}(S)$ and spend a large fraction of their time in \mathcal{T}_{ϵ_0} must have big curve complex distance at some point. We will repeatedly use Lemma 39 to “back up” along each of the two rays in order to make distance estimates.

Theorem 40 (Curve complex distance estimate). *Fix $0 < \delta < 1$. There is a constant $\lambda > 1$ such that for each $d > 6$ there exist constants $D_0, R_0 > 0$ with the following property. Given any $T \geq R_0$, let $\gamma_1 = [x, y]$ and $\gamma_2 = [x, z]$ be two Teichmüller geodesics based at x with lengths $r_1, r_2 \geq \lambda T$. Suppose that*

- (1) $\text{Th}_{\epsilon_0}^{\%}[x, y]$ and $\text{Th}_{\epsilon_0}^{\%}[x, z]$ are both at least δ ; and
- (2) for all $t \geq T$, the point $\gamma_1(t)$ is not contained in the D_0 -neighborhood of the geodesic $[x, z]$, and similarly for $\gamma_2(t)$ and $[x, y]$.

Then $d_S(y, z) \geq d$.

Proof. Choose λ such that $\lambda\delta > 1$ and define $\delta' = \frac{\lambda\delta-1}{\lambda-1}$; notice that $0 < \delta' < \delta$. Let ρ, R_1 and ρ', R'_1 be the corresponding constants guaranteed by Definite Progress (Proposition 38). Set k_0, M_0, N_0 , and c_0 to be the constants from the backing up lemma (Lemma 39) for our given d , and let $K \geq 1$ and $C \geq 0$ be large enough to be constants in the distance formula (1) for the threshold M_0 . Let D_0 and R_0 be any constants which satisfy

$$\begin{aligned} D_0 &> \max \left\{ N_0, \frac{2k_0d}{c_0\rho'}, \frac{2R'_1}{c_0}, K(2k_0+1)d + C \right\}, \text{ and} \\ R_0 &> \max \left\{ R_1, \frac{2d}{\rho} \right\}. \end{aligned}$$

Let $T \geq R_0$, $[x, y]$ and $[x, z]$ be as in the statement of the theorem; in order to derive a contradiction, assume furthermore that $d_S(y, z) < d$.

Set $y_0 = y$ and $z_0 = z$. By repeatedly backing up along γ_1 and γ_2 , we will recursively define sequences of points $\{y_i\}, \{z_i\}$. Each step of the sequence will cover a large Teichmüller distance and a comparatively small curve complex distance.

After backing up sufficiently far, we will eventually contradict the fact that each geodesic γ_i spends a large fraction of its time in \mathcal{T}_{ϵ_0} .

Suppose that the points y_i, z_i have been defined and satisfy

$$(\star) \quad d_S(y_i, z_i) < d; \quad d_{\mathcal{T}}(y_i, z_i) \geq D_0; \quad y_i, z_i \neq x.$$

(Notice that these conditions hold for the initial points y_0 and z_0 .) Backing up along the rays as in Lemma 39, we then obtain new points $y_{i+1} \in [x, y_i]$ and $z_{i+1} \in [x, z_i]$. For these, we have that $d_S(y_{i+1}, y_i)$ and $d_S(z_{i+1}, z_i)$ are bounded above by $k_0 \cdot d_S(y_i, z_i) < k_0 d$, by property (i) of backup points. We also have that either $[y_{i+1}, y_i]$ or $[z_{i+1}, z_i]$ (or both) has length at least $c_0 \cdot d_{\mathcal{T}}(y_i, z_i) \geq c_0 D_0$, by (iii). We are free to continue defining new points in this manner as long as the conditions of (\star) remain satisfied.

Suppose then that we have applied this procedure m times and arrived at points y_m and z_m . At each step of this process we traveled back a Teichmüller distance of at least $c_0 D_0$ along one of the two segments γ_1 or γ_2 . Therefore on at least one of the geodesics we have traveled a total Teichmüller distance of at least $c_0 D_0 m/2$. Without loss of generality, suppose γ_1 has this property; then $d_{\mathcal{T}}(y_m, y_0) \geq c_0 D_0 m/2 \geq R'_1$. On the other hand we have $d_S(y_m, y_0) \leq k_0 d m$, since at each step we travel at most $k_0 d$ in the curve complex. Therefore, along the geodesic segment $[y_m, y_0]$ the ratio of curve-complex distance to Teichmüller distance is

$$\frac{d_S(y_m, y_0)}{d_{\mathcal{T}}(y_m, y_0)} \leq \frac{2k_0 d m}{c_0 D_0 m} < \rho'.$$

Proposition 38 now implies that $\text{Th}_{\epsilon_0}^{\%}[y_m, y_0] < \delta'$.

Let $t_m \in [0, r_1]$ be the time for which $\gamma_1(t_m) = y_m$. We claim that if $t_m \geq r_1/\lambda$, then the points y_m, z_m satisfy (\star) so that we may reapply Lemma 39 and back up farther to points y_{m+1} and z_{m+1} . Firstly, since $t_m \geq r_1/\lambda \geq T$, hypothesis (ii) in the theorem implies that $y_m = \gamma_1(t_m)$ is not within D_0 of any point on γ_2 ; whence $d_{\mathcal{T}}(y_m, z_m) \geq D_0$. Applying the triangle inequality to the series of points $y_m, y_{m-1}, z_{m-1}, z_m$ implies that $d_S(y_m, z_m) \leq (2k_0 + 1)d$. Now if $d_Y(y_m, z_m) \leq M_0$ for all proper subsurfaces Y , then the distance formula (1) gives

$$d_{\mathcal{T}}(y_m, z_m) \leq K \cdot d_S(y_m, z_m) + C \leq K(2k_0 + 1)d + C < D_0,$$

which is a contradiction. Thus by Lemma 39 we have $d_S(y_m, z_m) \leq 6 < d$. Finally, since the fraction of $[t_m, r_1]$ that γ_1 spends in \mathcal{T}_{ϵ_0} is at most $\delta' < \delta$, it must be that the thick-stat along $[0, t_m]$ is at least δ (since the thick-stat of the whole of γ_1 is δ). By Proposition 38, since $t_m \geq T \geq R_1$, it now follows that $d_S(y_m, x) \geq \rho t_m \geq \rho R_0 \geq 2d$. As $d_S(y_m, z_m) \leq d$, we see that $z_m \neq x$. Thus we have verified the conditions of (\star) .

We have thus verified that we can repeatedly back up until we reach a point $y_m = \gamma_1(t)$ on $[x, y]$ (or similarly a point on $[x, z]$) such that $t < r_1/\lambda$ and the fraction of $[t, r_1]$ that γ_1 spends in \mathcal{T}_{ϵ_0} is strictly less than δ' . It now follows that the amount of time that γ_1 spends in \mathcal{T}_{ϵ_0} along $[0, r_1]$ is strictly less than

$$t + \delta'(r_1 - t) < \frac{r_1}{\lambda} + \frac{\lambda\delta - 1}{\lambda - 1} \left(\frac{\lambda - 1}{\lambda} r_1 \right) = \delta r_1,$$

again by shifting weight as in (10). This contradicts the first hypothesis of this theorem. \square

6.3. Teichmüller distance. The previous theorem implies that, under suitable conditions, we may assume our two Teichmüller geodesics stay far apart in the curve complex beyond some radius λT . We now show that in this situation, the distance between two points on the sphere of radius $r \gg \lambda T$ is on the order of $2r$.

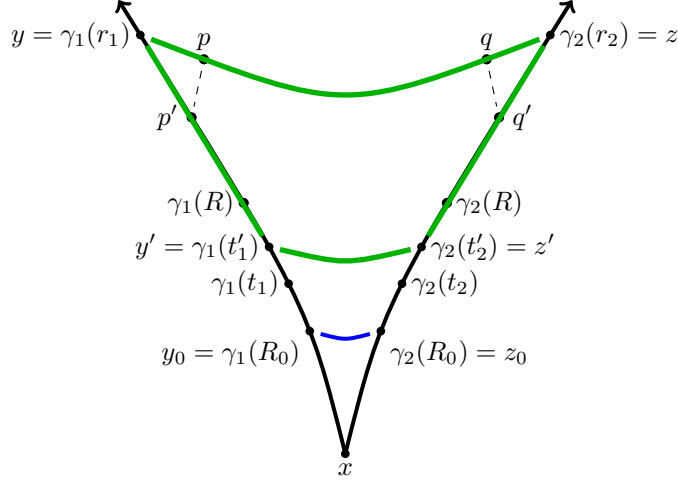


FIGURE 3. We will end up showing that $[y, z]$ “dips back” to within bounded distance of the basepoint and thus has length nearly $2r$.

Theorem 41 (Teichmüller distance estimate). *Fix any $0 < \delta < 1$. There exists a constant $H > 0$ such that for all sufficiently large R_0, d the following holds: If γ_1, γ_2 are geodesic rays based at $x \in \mathcal{T}(S)$ which satisfy*

- (a) $\text{Th}_{\epsilon_0}^{\%}[x, \gamma_i(r)] \geq \delta$ for all $r \geq R_0$ and $i = 1, 2$, and
- (b) for all $t_1, t_2 \geq R_0$ we have $d_S(\gamma_1(t_1), \gamma_2(t_2)) \geq d$,

then they also must satisfy

$$d_{\mathcal{T}}(\gamma_1(r_1), \gamma_2(r_2)) \geq r_1 + r_2 - HR_0$$

for all r_1, r_2 .

Proof. We fix some constants:

- Let $K \geq 1$ and $C \geq 0$ be large enough to be quasi-isometry constants in the distance formula (1) for threshold $10M$, and such that the projection of every geodesic in $\mathcal{T}(S)$ to $\mathcal{C}(S)$, concatenated with a geodesic segment of length at most 5, is an unparameterized (K, C) -quasigeodesic in $\mathcal{C}(S)$;
- Let τ be a constant such that (K, C) -quasigeodesic quadrilaterals in the curve complex are τ -thin;
- Let ρ, R_1 be the constants corresponding to δ from Proposition 38.
- We furthermore require that $R_0 > \max(R_1, 4 + C)$ and $d > 10 + \tau$. The desired constant H will not depend on R_0 or d .

Now write $y = \gamma_1(r_1)$, $y_0 = \gamma_1(R_0)$, $z = \gamma_2(r_2)$, $z_0 = \gamma_2(R_0)$, and let

$$\Omega = \{V \subsetneq S \mid T_V \cap [y_0, z_0] \neq \emptyset\}$$

be the set of subsurfaces which are thin somewhere along the geodesic segment $[y_0, z_0]$. For each $V \in \Omega$ it is possible that it determines a nonempty thin interval T_V which intersects $\gamma_i[R_0, \infty)$. We define

$$t_i = \sup \bigcup_{V \in \Omega} T_V \cap \gamma_i[R_0, \infty).$$

In other words, $t_i \in [R_0, \infty)$ is the smallest time such that for any $V \in \Omega$ and for $s > t_i$, $\gamma_i(s) \notin T_V$. It follows that for each $i = 1, 2$, there exists some $V_i \in \Omega$ such that the Bers marking at $\gamma_i(t_i)$ contains ∂V_i .

By definition of $V_i \in \Omega$, there exists a point $w_i \in [y_0, z_0]$ such that V_i is thin at w_i ; in particular $d_S(w_i, \gamma_i(t_i)) \leq 4$. Since w_i is at most R_0 away from an endpoint of $[y_0, z_0]$, the triangle inequality implies that $d_{\mathcal{T}}(x, w_i) \leq 2R_0$. We then have that

$$d_S(x, \gamma_i(t_i)) \leq 4 + d_S(x, w_i) \leq 4 + 2KR_0 + C < 3KR_0,$$

where the middle inequality is an application of the distance formula (1) and last inequality holds since $R_0 > 4 + C$. Furthermore, since $t_i \geq R_0$, our hypotheses imply that γ_i spends at least δt_i time during $[0, t_i]$ in \mathcal{T}_{ϵ_0} . Since $R_0 \geq R_1$, by Proposition 38 we have that

$$t_i \leq \frac{1}{\rho} d_S(x, \gamma_i(t_i)) < \frac{3K}{\rho} R_0.$$

Let $L = \frac{3KR_0}{\rho\delta}$, and note that $L > t_1, t_2$. If the geodesic segment $\gamma_i[t_i, L]$ were completely contained in the thin part, then the fraction of $[0, L]$ that γ_i spends in \mathcal{T}_{ϵ_0} would be at most

$$\frac{t_i}{L} < \frac{3KR_0}{\rho L} = \delta.$$

As this is not the case, there must exist a time $t'_i \in [t_i, L]$ at which $\gamma_i(t'_i) \in \mathcal{T}_{\epsilon_0}$. Let y' and z' equal $\gamma_i(t'_i)$ for $i = 1, 2$, respectively. Note that in particular

$$(12) \quad d_{\mathcal{T}}(x, y'), d_{\mathcal{T}}(x, z') \leq \frac{3KR_0}{\rho\delta}.$$

By the above estimate, the theorem is trivially true if either $r_i \leq t'_i$ if we choose $H \geq \frac{6K}{\rho\delta}$, so it is enough to prove the theorem for both $r_i > t'_i$. That is, $[x, y'] \subset [x, y]$ and $[x, z'] \subset [x, z]$. Now let

$$\Omega_1 = \{W \subsetneq S \mid d_W(y', y) \geq 10M\}, \quad \Omega_2 = \{W \subsetneq S \mid d_W(z', z) \geq 10M\}.$$

Note that a marking projects to the curve complex as a set with diameter 2. Also if $W_1 \in \Omega_1$ and $W_2 \in \Omega_2$ were disjoint or nested, then their boundaries would have distance ≤ 1 in the curve complex, so markings where those boundaries are short would have distance ≤ 5 . Thus in light of assumption (b) and since $d > 10$, we have that $W_1 \cap W_2$ for all $W_1 \in \Omega_1$ and $W_2 \in \Omega_2$. The same argument shows that $W_2 \in \Omega_2$ cannot determine an thin interval along $[y', y]$ and similarly for W_1 along $[z', z]$. In particular $\Omega_1 \cap \Omega_2 = \emptyset$.

Suppose V is any subsurface which is thin somewhere along $[y', y]$ (for example, if $V \in \Omega_1$). Since its thin interval T_V along $[x, y]$ is connected and y' is thick, we see that V is not thin along $[y_0, y']$. Furthermore, the definition of t_i implies that V is not thin along $[y_0, z_0]$, and assumption (b) plus the choice of d shows that V cannot be thin along $[z_0, z']$. Therefore, the triangle inequality gives $d_V(y', z') \leq 3M$ for these V . Letting

$$\Omega' = \{Z \subsetneq S \mid d_Z(y', z') \geq 10M\},$$

we conclude that each surface $Z \in \Omega'$ cannot be thin along $[y', y]$ (or $[z', z]$). In particular, the three collections Ω_1 , Ω_2 , and Ω' are pairwise disjoint.

Next we establish that for any $W \in \Omega_1 \cup \Omega_2 \cup \Omega'$,

$$d_W(y, z) \geq 6M.$$

This is because we know that for each $W \in \Omega_1$, W is not thin along $[z', z]$ and that $d_W(y', z') \leq 3M$ as above. Since $d_W(y', y) \geq 10M$, the triangle inequality gives $d_W(y, z) \geq 10M - 3M - M = 6M$. The same argument applies to Ω_2 . Similarly, for each $Z \in \Omega'$, we have both $d_Z(y', z') \geq 10M$ and $d_Z(y', y), d_Z(z', z) \leq M$, so that $d_Z(y, z) \geq 8M$. We conclude that each $W \in \Omega_1 \cup \Omega_2 \cup \Omega'$ determines a nonempty thin interval T_W along $[y, z]$.

Let $p \in [y, z]$ denote the last point (in traveling from y to z) at which any subsurface in Ω_1 is thin and let $W_1 \in \Omega_1$ denote the corresponding subsurface. Analogously define q and $W_2 \in \Omega_2$ on $[z, y]$; p' and $Z_1 \in \Omega_1$ on $[y, y']$; and q' and $Z_2 \in \Omega_2$ on $[z, z']$. (If $\Omega_i = \emptyset$ then we take the initial endpoints of the respective intervals and leave the associated subsurfaces undefined.) Thus the Bers markings at p, q, p', q' contain $\partial W_1, \partial W_2, \partial Z_1, \partial Z_2$, respectively.

It cannot be the case that $W_1 \cap Z_1$, as this would violate the time-order principle (W_1 occurs after Z_1 on $[y, z]$ but before Z_1 on $[y, y']$). Therefore ∂W_1 and ∂Z_1 either coincide or are disjoint; in either case we have $d_S(p, p') \leq 5$. The same argument shows that $d_S(q, q') \leq 5$.

Next we want to verify that p occurs before q along $[y, z]$. This is another application of time ordering: suppose for contradiction that T_{W_2} appeared before T_{W_1} along $[y, z]$. (Recall that $W_1 \cap W_2$, since this was verified for all subsurfaces from Ω_1 and Ω_2 above.) Then since $d_{W_1}(y, z), d_{W_2}(y, z), d_{W_1}(y', y)$ are all more than $3M$, Lemma 11 ensures that T_{W_2} also appears along $[y, y']$. But this contradicts the fact that no surface in Ω_2 has a thin interval along $[y, y']$.

The setup is now complete, and we are ready to prove the following:

Claim 42. *There exists a constant $\lambda > 0$ (independent of r_1, r_2, R_0, d) such that the minimal distance between the segments $[y, z]$ and $[y', z']$ is at most λR_0 .*

It is easy to see that the theorem follows from this claim. Suppose that the geodesic segment $[y, z]$ comes within distance λR_0 of the geodesic $[y', z']$. As this latter segment is contained in the ball of radius $2L = \frac{6K}{\rho\delta} R_0$ around x , this means that there must be some point w on $[y, z]$ whose distance from x satisfies $d_{\mathcal{T}}(w, x) \leq \frac{6K}{\rho\delta} R_0 + \lambda R_0$. Since y and z are on the spheres of radius r_1, r_2 centered at x , this proves that

$$d_{\mathcal{T}}(y, z) = d_{\mathcal{T}}(y, w) + d_{\mathcal{T}}(w, z) \geq r_1 + r_2 - 2 \left(\frac{6K}{\rho\delta} + \lambda \right) R_0,$$

establishing the theorem with $H = 2 \left(\frac{6K}{\rho\delta} + \lambda \right)$.

Proof of claim. Consider any two points $w' \in [y', z']$ and $w \in [p, q]$. Aiming to use the distance formula (1), we bound the projections $d_V(w, w')$ in terms of $d_V(y', z')$ for all proper subsurfaces $V \subsetneq S$.

As usual, it will require special care to deal with the case when V is an annulus for which the reverse triangle inequality fails. Let Υ denote the set of annular subsurfaces V for which $w \in T_V$ and Lemma 12 fails for V along $[y, z]$. Let Υ'

denote the analogous set for w' and $[y', z']$. Notice that, since at most $3g - 3$ curves can be short at w or w' , we have $|\Upsilon|, |\Upsilon'| \leq 3g - 3$.

Below, it will be essential to bound the numbers $d_V(y', w')$, $d_V(w', z')$, $d_V(y, w)$, and $d_V(w, z)$. First, suppose that $V \notin \Upsilon$ because $w \notin T_V$. Then, by the connectedness of T_V , we have either $d_V(y, w) \leq M$ or $d_V(w, z) \leq M$; whence, by the triangle inequality, we may conclude that

$$d_V(y, w), d_V(w, z) \leq d_V(y, z) + M.$$

On the other hand, the reverse triangle inequality implies that the same bounds hold for every other $V \notin \Upsilon$ as well. For all $V \notin \Upsilon'$ we have the analogous bounds

$$d_V(y', w'), d_V(w', z') \leq d_V(y', z') + M.$$

We now proceed to bound $d_V(w', w)$ for all proper subsurfaces. Firstly, any $V \in \Omega_1$ is not thin along $[w, z]$ or $[z, z']$; therefore,

$$d_V(w', w) \leq d_V(w', z') + d_V(z', z) + d_V(z, w) \leq d_V(w', z') + 2M$$

Similarly, any $V \in \Omega_2$ is neither thin on $[y', y]$ nor $[y, w]$, so we have

$$d_V(w', w) \leq d_V(w', y') + d_V(y', y) + d_V(y, w) \leq d_V(w', y') + 2M.$$

In either of the above cases, these bounds reduce to

$$d_V(w', w) \leq d_V(y', z') + 3M$$

provided that $V \notin \Upsilon'$. If V is in Υ' , applying Theorem 13 with constant $35M$, there exists B' such that either

$$d_V(y', w'), d_V(w', z') \leq d_V(y', w') + d_V(w', z') \leq d_V(y', z') + B',$$

or there is a collection $\{U_j\}$ of subsurfaces that do satisfy Lemma 12 along $[y', z']$, are disjoint from V , each satisfy $d_{U_j}(y', w') \geq 35M$, and such that

$$\begin{aligned} d_V(y', w') &\leq \sum_{U_j} d_{U_j}(y', w') \leq \sum_{U_j} (d_{U_j}(y', z') + B) \leq \sum_{U_j} 2d_{U_j}(y', z'), \\ &\leq 2 \sum_{Y \subsetneq S} [d_Y(y', z')]_{10M} \leq 2Kd_{\mathcal{T}}(y', z') + 2C, \end{aligned}$$

where the second inequality above follows from Lemma 12, and the third and fourth from the fact that $d_{U_j}(y', z') \geq d_{U_j}(y', w') - B \geq 34M \geq B$. Replacing y' with z' , we obtain the same bound on $d_V(z', w')$ as well. (Notice that these bounds on $d_V(y', w')$ and $d_V(z', w')$ do not use the assumption $V \in \Omega_1 \cup \Omega_2$.) To summarize, for all $V \in \Omega_1 \cup \Omega_2$, we have that

$$d_V(w', w) \leq \begin{cases} d_V(y', z') + 3M, & \text{if } V \notin \Upsilon' \\ d_V(y', z') + 3M + B' + 2Kd_{\mathcal{T}}(y', z') + 2C, & \text{if } V \in \Upsilon'. \end{cases}$$

Lastly, for a proper subsurface $V \notin \Omega_1 \cup \Omega_2$ (such as $V \in \Omega'$), we have

$$d_V(w', w) \leq d_V(w', y') + d_V(y', y) + d_V(y, w) \leq d_V(y', w') + 10M + d_V(y, w).$$

The same holds with y, y' replaced with z, z' . As above, the $d_V(w', y')$ (or $d_V(w', z')$) term may be bounded by either

$$d_V(y', z') + M \quad \text{or} \quad d_V(y', z') + M + B' + 2Kd_{\mathcal{T}}(y', z') + 2C$$

depending on whether or not $V \notin \Upsilon'$. If $V \notin \Upsilon$, the $d_V(y, w)$ (or $d_V(w, z)$) term is bounded similarly:

$$d_V(y, w), d_V(w, z) \leq d_V(y, z) + \mathbf{M} \leq d_V(y', z') + 21\mathbf{M}.$$

When $V \in \Upsilon$, we instead bound the $d_V(y, w)$ (or $d_V(w, z)$) term as follows: By Theorem 13 (applied with constants $35\mathbf{M}$ and B'), we either have

$$\begin{aligned} d_V(y, w), d_V(w, z) &\leq d_V(y, w) + d_V(w, z) \leq d_V(y, z) + B' \\ &\leq d_V(y', z') + 20\mathbf{M} + B', \end{aligned}$$

or there exists a collection $\{U_j\}$ of subsurfaces that are disjoint from V , satisfy Lemma 12 and $d_{U_j}(y, w) \geq 35\mathbf{M}$, and such that

$$d_V(y, w) \leq \sum_{U_j} d_{U_j}(y, w).$$

Lemma 12 implies that $d_{U_j}(y, z) \geq 34\mathbf{M}$ so that each U_j has large projections along both $[y, w]$ and $[y, z]$. The only such subsurfaces are contained in $\Omega_1 \cup \Omega'$ (since $W \notin \Omega_1 \cup \Omega' \cup \Omega_2$ implies $d_W(y, z) \leq 30\mathbf{M}$). Running the argument with y replaced by z , we obtain another such collection contained in $\Omega_2 \cup \Omega'$. Since all of these subsurfaces are disjoint from V , the fact that elements of Ω_1 and Ω_2 are far apart in $\mathcal{C}(S)$ implies that one of these collections must, in fact, be contained in Ω' . Therefore we may assume that $\{U_j\} \subset \Omega'$. We then have

$$d_{U_j}(y, w) \leq d_{U_j}(z, y) + B \leq d_{U_j}(y', z') + 20\mathbf{M} + B \leq 3d_{U_j}(y', z'),$$

where the last inequality holds because $d_{U_j}(y', z') \geq 14\mathbf{M}$. Thus we conclude that

$$d_V(y, w) \leq \sum_{U_j} d_{U_j}(y, w) \leq 3 \sum_{Y \subsetneq S} [d_Y(y', z')]_{10\mathbf{M}} \leq 3Kd_{\mathcal{T}}(y', z') + 3C.$$

Putting these estimates together, we see that for all $V \notin \Omega_1 \cup \Omega_2$ we have

$$d_V(w', w) \leq \begin{cases} 2d_V(y', z') + 32\mathbf{M}, & \text{if } V \notin \Upsilon' \cup \Upsilon \\ 2d_V(y', z') + 32\mathbf{M} + 2B' + 5Kd_{\mathcal{T}}(y', z') + 5C, & \text{if } V \in \Upsilon' \cup \Upsilon. \end{cases}$$

In fact, comparing with the above estimates for $W \in \Omega_1 \cup \Omega_2$, we see that these bounds hold for *all* proper subsurfaces $V \subsetneq S$.

Observe that for any threshold $N > 0$ we have

$$[A + B]_N \leq 2[A]_{N/2} + 2[B]_{N/2} \quad \text{and} \quad [2A]_N = 2[A]_{N/2}$$

for all $A, B \geq 0$. Set a threshold $N = 200\mathbf{M} + 10B' + 25C$. Then, for all $V \subsetneq S$,

$$\begin{aligned} [d_V(w, w')]_N &\leq [2d_V(y', z') + 32\mathbf{M}]_N \leq 2[2d_V(y', z')]_{N/2} + 2[32\mathbf{M}]_{N/2} \\ &= 4[d_V(y', z')]_{N/4} \leq 8[d_V(y', z')]_{N/8} \end{aligned}$$

provided that $V \notin \Upsilon' \cup \Upsilon$. We similarly have

$$\begin{aligned} [d_V(w, w')]_N &\leq [2d_V(y', z') + 32\mathbf{M} + 5Kd_{\mathcal{T}}(y', z') + 5C]_N \\ &\leq 4[2d_V(y', z')]_{N/4} + 4[32\mathbf{M}]_{N/4} + [5Kd_{\mathcal{T}}(y', z')]_{N/4} + 4[2B' + 5C]_{N/4} \\ &\leq 8[d_V(y', z')]_{N/8} + 5Kd_{\mathcal{T}}(y', z') \end{aligned}$$

when $V \in \Upsilon' \cup \Upsilon$. Moreover, recall from (12) that $d_{\mathcal{T}}(y', z') \leq \frac{6K}{\rho\delta}R_0$. Letting K' and C' be sufficiently large quasi-isometry constants such that the distance formula

(1) holds for the thresholds N and $N/8$ (note that K', C' are independent of R_0), we conclude that

$$\begin{aligned}
 d_{\mathcal{T}}(w, w') &\leq K' \left(d_S(w, w') + \sum_{V \subsetneq S} [d_V(w, w')]_N \right) + C' \\
 &\leq K' \left(d_S(w, w') + \sum_{V \subsetneq S} 8 [d_V(y', z')]_{N/8} + \sum_{V \in \Upsilon' \cup \Upsilon} 5K d_T(y', z') \right) + C' \\
 &\leq K' \left(d_S(w, w') + 8K' d_{\mathcal{T}}(y', z') + 8C' + (6g - 6)5K d_T(y', z') \right) + C' \\
 &\leq K' d_S(w, w') + K'(8K' + 30gK) \frac{6K}{\rho\delta} R_0 + 9K' C'
 \end{aligned}$$

Thus we have bounded the Teichmüller distance across the green quadrilateral in Figure 3 linearly in terms of the corresponding curve complex distance and R_0 . We will now use the fact that quasigeodesic quadrilaterals in the curve complex are thin. Let π be the projection map to the curve complex. On the left, we let γ_L be the concatenation of the quasi-geodesic segment $\pi[y', p']$ with the geodesic $[\pi(p'), \pi(p)]$; likewise, on the right, γ_R is the concatenation of $\pi[z', q']$ and $[\pi(q'), \pi(q)]$. These are actually (K, C) -quasigeodesics because the second segment has bounded length: $d_S(p, p'), d_S(q, q') \leq 5$ (recall the definition of K and C). Thus the quasi-geodesic segments γ_L , $\pi[p, q]$, γ_R , and $\pi[y', z']$ form a (K, C) -quasigeodesic quadrilateral, and we conclude that each side is contained in the τ -neighborhood of the union of the other three sides.

Now, the separation hypothesis (b) of the theorem implies that no point on $\pi[y', p']$ is within d of any point on $\pi[z', q']$. Therefore, no point on γ_L is within $(d - 10)$ of any point on γ_R . Recall that $d - 10 > \tau$, which implies that no point on γ_L is contained in the τ -neighborhood of γ_R . This means that there exist points on $\pi[p, q]$ and $\pi[y', z']$ that are within $2\tau + 1$ of each other (since each point on γ_L , say, is within τ of one or the other, and the points of γ_L are separated by one). By definition of projection distance, these correspond to points $w \in [p, q], w' \in [y', z']$ with $d_S(w, w') \leq 2\tau + 1$. Combining this with our above estimates and using the fact $R_0 \geq 1$, we find that

$$d_{\mathcal{T}}(w, w') \leq \left(K'(2\tau + 1) + (8K' + 30gK) \frac{6K'K}{\rho\delta} + 9K' C' \right) R_0$$

This completes the proof of Claim 42 and Theorem 41. \square

7. STATISTICAL HYPERBOLICITY

We can now assemble our results to prove Theorems 1 and 2. Our distance estimates deal with annuli of fixed width.

Theorem 43. *Fix a basepoint $x \in \mathcal{T}(S)$ and let μ be any measure on $\mathcal{T}(S)$ satisfying a thickness estimate and an exponential decay estimate. Fix an arbitrary $k > 0$. Then*

$$\lim_{r \rightarrow \infty} \frac{1}{r} \frac{1}{\mu(\mathcal{A}_r^k(x))^2} \int_{\mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x)} d_{\mathcal{T}}(y, z) d\mu(y) d\mu(z) = 2.$$

We have shown that these hypotheses are satisfied by the standard visual measures $\text{Vis}(\nu_x)$ and $\text{Vis}(s_x)$ (Proposition 24, Theorem 27), as well as the holonomy measure \mathbf{m} (Theorem 25, Theorem 28), and therefore the Hausdorff measure η , the symplectic measure \mathbf{n} , and the Finsler measures considered above.

Proof. Set $\delta = 1/2$ and let λ, H be the constants guaranteed by Theorems 40 and 41. Choose constants $\epsilon > 0$, $\sigma > 0$, and $M_0 > 0$. After choosing a sufficiently large thick part \mathcal{T}_{ϵ_0} and restricting to a subset $E_r \subset \mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x)$ whose complement has proportional measure at most ϵ (for sufficiently large r), we may suppose that all pairs $(y_1, y_2) \in E_r$ satisfy

$$\text{Th}_{\epsilon_0}^{\%}[x, \gamma_i(t)] \geq \delta \quad \text{and} \quad d_{\mathcal{T}}(\gamma_1(t), \gamma_2(t)) \geq M_0$$

for all $t \geq \sigma r$, where γ_i is the geodesic ray based at x through y_i . Notice that, in this case, the point $\gamma_1(t)$ cannot be within $M_0/2$ of any point on the geodesic ray γ_2 (and similarly for $\gamma_2(t)$ and γ_1). Therefore, since M_0 may be chosen arbitrarily large, for each $d > 6$ Theorem 40 guarantees that $d_S(\gamma_1(t_1), \gamma_2(t_2)) \geq d$ for all $r \gg 1$ and $t_1, t_2 \geq \lambda \sigma r$. Choosing d to be sufficiently large, we may then apply Theorem 41 to conclude that for all $r \gg 1$ and all $(y_1, y_2) \in E_r$ we have

$$d_{\mathcal{T}}(\gamma_1(r_1), \gamma_2(r_2)) \geq r_1 + r_2 - H\lambda\sigma r$$

for all r_1, r_2 . Putting the above estimates together, we find that

$$\begin{aligned} \liminf_{r \rightarrow \infty} \frac{1}{r} \frac{1}{\mu(\mathcal{A}_r^k(x))^2} \int_{\mathcal{A}_r^k(x) \times \mathcal{A}_r^k(x)} d_{\mathcal{T}}(y, z) d\mu(y) d\mu(z) \\ \geq \liminf_{r \rightarrow \infty} \frac{1}{r} (1 - \epsilon)(2r - 2k - H\lambda\sigma r) = (1 - \epsilon)(2 - H\lambda\sigma). \end{aligned}$$

Since ϵ and σ can be chosen arbitrarily small, the result follows. \square

We immediately deduce the same result for balls if the measure has definite exponential growth.

Theorem 1. *Fix a basepoint x and let μ be any measure satisfying a thickness estimate and an exponential decay estimate, for which there is definite exponential growth $C_1 e^{\alpha r} \leq \mu(\mathcal{B}_r(x)) \leq C_2 e^{hr}$. Then*

$$\lim_{r \rightarrow \infty} \frac{1}{r} \frac{1}{\mu(\mathcal{B}_r(x))^2} \int_{\mathcal{B}_r(x) \times \mathcal{B}_r(x)} d_{\mathcal{T}}(y, z) d\mu(y) d\mu(z) = 2.$$

The growth condition, verified for our measures in Corollary 20, ensures that most of the ball is concentrated in the annulus (Lemma 21).

Theorem 2. *Fix $x \in \mathcal{T}(S)$. For either of the standard visual measures $\mu_x = \text{Vis}_r(\nu_x)$ or $\text{Vis}_r(s_x)$ on the Teichmüller sphere $S_r(x)$ of radius r , we have that*

$$E(\mathcal{T}(S), \mu_x) = 2.$$

Proof. Visual measures on $\mathcal{T}(S)$ are constructed by radially integrating these visual measures on spheres; in fact, Proposition 24 and Theorem 27 from §5 were proved for annuli by first verifying them for spheres. Thus we have both a thickness estimate and an exponential decay estimate for visual measures on spheres. The result follows. \square

REFERENCES

- [1] J.C. Álvarez Paiva, A. Thompson, *Volumes on normed spaces and Finsler spaces*, MSRI Publications 50 (2004).
- [2] J.S. Athreya, Quantitative recurrence and large deviations for Teichmüller geodesic flow. *Geom. Dedicata* 119 (2006), 121–140.
- [3] J. Athreya, A. Bufetov, A. Eskin, and M. Mirzakhani, *Lattice point asymptotics and volume growth on Teichmüller space*, arxiv 0610715.
- [4] L. A. Santaló, *An affine invariant for convex bodies of n -dimensional space*, *Portugaliae Math.* 8 (1949), 155–161.
- [5] H. Busemann, *Intrinsic area*, *Ann. of Math.* 48 (1947), 234–267.
- [6] M. Duchin, S. Lelièvre, and C. Mooney, *Statistical hyperbolicity in groups*, *Alg. and Geom. Top.* 12 (2012), 1–18.
- [7] A. Eskin and M. Mirzakhani, *Counting closed geodesics in moduli space*, arxiv 0811.2362.
- [8] A. Fathi, F. Laudenbach, P. Poenaru, *Travaux de Thurston sur les surfaces*, *Asterisque* 66-67 (1979).
- [9] F. John, *Extremum problems with inequalities as subsidiary conditions*, 187–204 in *Studies and essays presented to R. Courant on his 60th birthday*, edited by K. O. Friedrichs, O. Neugebauer, and J. J. Stoker, Interscience, New York, 1948.
- [10] H. Masur, *On a class of geodesics in Teichmüller space*, *Ann. of Math.* (2) 102 (1975), no. 2, 205–221.
- [11] H. Masur, *Uniquely ergodic quadratic differentials*, *Comment. Math. Helvetici* 55 255–266 (1980).
- [12] H. Masur, *Interval exchange transformations and measured foliations*, *Annals of Math.* 115 (1982) 169–200.
- [13] H. Masur, *Hausdorff dimension of the set of nonergodic foliations of a quadratic differential*, *Duke Math. Journal* 66 387–442 (1992).
- [14] H. Masur, *The Teichmüller flow is Hamiltonian*, *Proceedings of the American Mathematical Society* 123 (1995) 12 3739–3747.
- [15] H. Masur, and Y. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, *Invent. Math.* 138 (1999), no. 1, 103–149.
- [16] H. Masur and M. Wolf, *Teichmüller space is not Gromov hyperbolic*, *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 20(2) (1995), 259–267.
- [17] Y.N. Minsky, *Extremal length estimates and product regions in Teichmüller space*, *Duke Math. J.* 83:2(1996), 249–286.
- [18] K. Rafi, *A combinatorial model for the Teichmüller metric*, *Geom. Funct. Anal.* 17:3 (2007), 936–959.
- [19] K. Rafi, *Hyperbolicity in Teichmüller space*, arxiv 1011.6004.
- [20] K. Rafi and S. Schleimer, *Covers and the curve complex*, *Geom. Topol.* 13 (2009), no. 4, 2141–2162.
- [21] H. L. Royden, *Automorphisms and isometries of Teichmüller space*, *Advances in the Theory of Riemann Surfaces* (Proc. Conf., Stony Brook, N.Y., 1969), *Ann. of Math. Studies*, No. 66. Princeton Univ. Press, Princeton, N.J., (1971), 369–383.

SPENCER DOWDALL, MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN,
 URBANA, IL 61801
E-mail address: `dowdall@math.uiuc.edu`

MOON DUCHIN, MATHEMATICS, TUFTS UNIVERSITY, MEDFORD, MA 02155
E-mail address: `moon.duchin@tufts.edu`

HOWARD MASUR, MATHEMATICS, UNIVERSITY OF CHICAGO, CHICAGO, IL 60637
E-mail address: `masur@math.uchicago.edu`